

UNIVERSIDAD DE LOS ANDES
FACULTAD DE CIENCIAS ECONÓMICAS Y SOCIALES
INSTITUTO DE ESTADÍSTICA APLICADA Y COMPUTACIÓN
PROGRAMA DE DOCTORADO EN ESTADÍSTICA

ESTIMACIÓN DE PARÁMETROS PARA EL MODELO LINEAL EN p ETAPAS.
PERSPECTIVA DESDE EL PUNTO DE VISTA DE LOS OPERADORES DE
PROYECCIÓN

www.bdigital.ula.ve

Ing. Jorge Luis Valera Briceño
Tutor: Doctor Surendra Prasad Sinha
Co-Tutor: Doctor Arnaldo Goitía

TESIS DOCTORAL

Presentada ante la ilustre Universidad de Los Andes para optar al
Grado Académico de Doctor en Estadística.

MÉRIDA, VENEZUELA

Octubre, 2013

www.bdigital.ula.ve

Derechos reservados
Jorge Luis Valera Briceño
(2013)

Reconocimiento-No comercial-Compartir igual

Dedicatoria

Dedico este triunfo a Dios todopoderoso, fuente de vida y apoyo incondicional en el logro de mis metas.

A mi madre, quien siempre ha estado a mi lado en todo momento. Gracias mamá por tu amor, tu comprensión, y tus oraciones en todo momento y a lo largo de mi vida. Te amo mamá.

A mis hermanos, Xiomara, Gerardo, Noel, y muy especialmente a mi hermano Lino por todo su apoyo al abrirme las puertas de su hogar.

A mi esposa Carolina, por su amor, su paciencia y comprensión.

A mi hija Isabella María, porque sus sonrisas y sus alegrías son mi mayor inspiración para alcanzar este triunfo y seguir adelante. Eres mi vida.

A mi hijo Jorge Luis, que está por nacer y junto a Isabella son parte fundamental de mis alegrías de cada día.

Jorge Luis Valera Briceño

Agradecimientos

Deseo expresar mi agradecimiento a todas aquellas personas que me dieron su apoyo en diversas formas. gracias al cual he logrado alcanzar esta meta. Especialmente a

A la Ilustre Universidad de los Andes, por darme la oportunidad para el logro de esta meta.

Al Doctor Surendra Prasad Sinha, mi tutor, por la confianza puesta en mí al aceptar ser mi guía en esta investigación. Por su disposición, paciencia y tiempo al orientarme hacia la culminación con éxito de este trabajo.

Al Doctor Ernesto Ponsot Balaguer, por su colaboración y recomendaciones que permitieron mejorar esta investigación.

A las Señoras Lucy y Clemencia, por su excelente labor en el desempeño de sus funciones en el I.E.A.C., por su atención y cortesía.

A todos los profesores del I.E.A.C que han sido parte de mi formación académica.

Resumen

ESTIMACIÓN DE PARÁMETROS PARA EL MODELO LINEAL EN p ETAPAS.
PERSPECTIVA DESDE EL PUNTO DE VISTA DE LOS OPERADORES DE
PROYECCIÓN

por

Jorge Luis Valera Briceño

Este trabajo tiene como objetivo fundamental, obtener expresiones alternativas para los estimadores del vector de medias y del vector de parámetros del modelo lineal en dos etapas, desde el enfoque de los operadores de proyección, y generalizar estos resultados al modelo lineal en p etapas regular. Para ello, inicialmente se llevó a cabo una investigación documental relacionada con la teoría de la estimación, los operadores de proyección de espacios vectoriales, entre otros. Se analizaron los operadores de proyección y las funciones lineales cero desde el punto de vista de los modelos lineales, y se estudiaron los modelos en etapas. Utilizando la expresión para el estimador del vector de medias de un modelo lineal dada por Bhimasankaran y Sengupta (la cual es un caso particular de la expresión dada por Rao) junto con la condición para obtener estimadores UBLUE en un modelo en dos etapas dada por Volaufova, se obtuvieron expresiones para los estimadores del vector de parámetros y del vector de medias para este modelo, en la forma de operadores de proyección. Luego los resultados obtenidos fueron generalizados al caso de un modelo con p etapas.

Palabras Claves: Estimación, Operadores de proyección, Funciones Lineal Cero, UBLUE, Proyectores, Modelos lineales en etapas.

Índice general

Introducción	1
1. Planteamiento del Problema y Objetivos	3
1.1. Planteamiento del Problema	3
1.2. Objetivos	4
1.2.1. Objetivo General	5
1.2.2. Objetivos Específicos	5
1.3. Antecedentes e Importancia de la Investigación	5
2. Marco Teórico	13
2.1. Definiciones del álgebra lineal	13
2.1.1. Operadores de proyección de un espacio vectorial	14
2.2. Los Modelos Lineales	16
2.3. Estimación	17
2.3.1. Estimación por Máxima Verosimilitud	17
2.3.2. Estimadores UMVU para β y σ^2	20
2.3.3. Estimadores de Mínimos Cuadrados	21
2.4. Identificabilidad, estimabilidad y estimadores BLUE	23
2.4.1. Estimadores de Mínimos Cuadrados Generalizados	26
2.5. Estimadores UBLUE	28
2.6. Los operadores de proyección y los BLUE	28
2.7. Las funciones lineales cero (LZF).	33
2.7.1. Definición	33

2.7.2. Relación de las funciones lineales cero con las funciones lineales estimables	33
3. Marco Metodológico	35
3.1. Tipos de Investigación	35
3.1.1. Investigación Documental	35
3.1.2. Investigación Exploratoria	36
3.1.3. Investigación Descriptiva	36
3.2. Diseño de la Investigación	36
4. El Modelo Lineal en Dos Etapas Regular.	38
4.1. Resultados Preliminares	38
4.1.1. Algunos Resultados de Interés.	39
4.1.2. Caracterización del BLUE de $X\beta$	42
4.1.3. Las funciones lineales cero y el BLUE de $X\beta$	45
4.1.4. El estimador de $X\beta$ cuando V es no negativa definida	47
4.2. Los modelos lineales en dos etapas	51
4.3. UBLUE para $X^*\beta$ cuando $\rho = \sigma_2^2/\sigma_1^2$ es conocido	53
4.4. Los UBLUE de β y $X\beta$ cuando $\rho = \sigma_2^2/\sigma_1^2$ es conocido	58
4.5. Los UBLUE de β y $X\beta$ para todo $\sigma_1^2 > 0, \sigma_2^2 > 0$	59
5. Generalización al modelo de p-etapas	65
5.1. Relación entre el modelo lineal regular y el modelo transformado	65
5.1.1. Caso: Modelo en dos etapas	65
5.1.2. Caso: Modelo en tres etapas	67
5.1.3. Caso: Modelo en cuatro etapas	69
5.1.4. Caso: Modelo en cinco etapas	71
5.2. Procedimiento iterativo para obtener F	73
5.2.1. Equivalencia de las expresiones para obtener F_p	81
5.3. Obtención del estimador UBLUE para β y $X\beta$	91

6. Aplicación de un modelo en dos etapas	99
6.1. El problema del Rendimiento Estudiantil	101
6.1.1. Aplicación del modelo en dos Etapas Regular	102
6.1.2. Población y muestra	103
6.2. Ajuste del modelo en dos etapas	105
6.2.1. Análisis descriptivo y ajuste de los datos a una distribución normal	105
6.3. Ajuste del modelo para los datos de la muestra	107
6.4. Ajuste del modelo para varias muestras	111
6.5. Conclusiones	114
Bibliografía	115

www.bdigital.ula.ve

Índice de cuadros

6.1. Correlaciones entre las variables bajo estudio.	105
6.2. Resumen estadístico de las variables.	107
6.3. Ajuste de las variables respuesta a una distribución normal.	107
6.4. Resultados de ajustar el modelo en dos etapas.	109
6.5. Promedios, mínimos y máximos de los parámetros por grupo.	110
6.6. Intervalos de confianza para los parámetros del modelo por grupo de muestras.	112

Índice de figuras

6.1. Fuente: Elaboración propia. Correlaciones entre las variables de la muestra	106
6.2. Valores observados y predichos a partir del modelo ajustado. Fuente: Elaboración propia.	110
6.3. Error = Valores observados - Valores predichos. Fuente: Elaboración propia.	111
6.4. Intervalos de confianza para los parámetros del modelo en dos etapas. Fuente: Elaboración propia.	113

Introducción

El estudio de los modelos lineales es una de las ramas de la estadística de mayor interés para los investigadores, dada su gran utilidad en el análisis de datos, en diversas áreas del saber. La gran cantidad de literatura relacionada con los modelos lineales, y las investigaciones teóricas y prácticas llevadas a cabo, avalan la importancia que este tipo de modelos tienen para los investigadores.

Según Graybill (1976) un modelo lineal general es $y = X\beta + \varepsilon$; en donde y es un vector aleatorio observable, X una matriz real observable, β un vector de parámetros desconocidos, y ε un vector aleatorio no observable tal que $E(\varepsilon) = 0$ y $Cov(\varepsilon) = \Sigma$. Los vectores y , β , ε y la matriz X de dimensiones congruentes.

Cuando se estudian los modelos lineales, el interés recae fundamentalmente en la estimación de los parámetros del modelo y de algunas funciones lineales que dependen de estos parámetros, a las cuales se les conoce como funciones estimables. Dependiendo de la distribución del vector de error ε (o bien del vector de respuestas y) y de las propiedades lineales de la matriz X , se determina el conjunto de las funciones lineales del vector de parámetros que es posible estimar y sus estimadores, como funciones lineales del vector de respuestas y .

INTRODUCCIÓN

El estimador de $X\beta$, obtenido del método de mínimos cuadrados o a través del método de máxima verosimilitud, es el que se utiliza para predecir al vector de respuestas y . Cuando la matriz $\Sigma = \sigma^2 V$, el estimador de $X\beta$ obtenido por cualquiera de los métodos mencionados posee propiedades que lo convierten en el mejor estimador lineal insesgado conocido como el BLUE de $X\beta$, el cual es también, el estimador insesgado de varianza mínima uniformemente (UMVUE) si la distribución del vector ε (o de y) es normal multivariante.

Dentro de este tipo de modelos se encuentran los modelos lineales en etapas, los cuales han sido estudiados por Kubáček y Volafova en varios artículos. En este estudio, primero se obtienen expresiones para los estimadores del vector de parámetros $\beta = (\beta'_1, \beta'_2)'$, y del vector de medias $X\beta$ para un modelo en dos etapas desde el punto de vista de la metodología de los operadores de proyección, y luego estos resultados son extendidos al caso de los modelos en p etapas. Para el logro de este objetivo se emplean fundamentalmente los resultados dados por Bhimasankaran y Sengupta (1996) y por Rao (1974).

Este trabajo consta de seis capítulos. En el capítulo 1 se expone el planteamiento del problema, los objetivos y los antecedentes e importancia de la investigación. El capítulo 2 presenta el marco teórico. Aquí se dan los desarrollos teóricos básicos necesarios para llevar a cabo este estudio. El marco metodológico se presenta en el capítulo 3 y en el se describen los aspectos relacionados con el desarrollo de la investigación a fin de exponer el tipo de estudio que se realizó. En el capítulo 4 se estudia el modelo en dos etapas y se dan resultados relacionados con este modelo. Posteriormente en el capítulo 5 los resultados obtenidos en el capítulo 4 son generalizados al caso de un modelo en p etapas y finalmente en el capítulo 6 se da un ejemplo de aplicación, en donde se utiliza un modelo en dos etapas regular.

Capítulo 1

Planteamiento del Problema y Objetivos

1.1. Planteamiento del Problema

Muchas investigaciones emplean modelos lineales con el fin explicar el comportamiento de una variable respuesta o dependiente en función de una o más variables explicativas o independientes. Para ello, se ajustan modelos lineales con el fin de realizar pronósticos relacionados con la variable respuesta, o para hacer inferencia sobre su valor promedio o esperado, en función de valores obtenidos de las variables explicativas. Para obtener buenas predicciones e inferencias relacionadas con la variable respuesta, en la teoría de los modelos lineales se han desarrollado métodos eficientes para obtener buenos estimadores de los parámetros del modelo lineal ajustado, los cuales consideran las características particulares de cada modelo.

Con frecuencia los estudios que se realizan para comprender el comportamiento de las variables dependientes no son aislados, sino por lo contrario tienden a llevarse a cabo repetidamente, como es el caso de las investigaciones relacionadas con el mejoramiento de un producto, en las cuales la variable respuesta de interés es observada cada vez que las condiciones de elaboración del producto cambian a lo largo de todo el proceso de mejoramiento.

Bajo estas circunstancias, cada vez que se realiza nueva experimentación para es-

tudiar determinada variable respuesta, las investigaciones previas relacionadas con ella deberían poder aprovecharse para brindar información que permita mejorar las estimaciones o predicciones, en la investigación actual. En este sentido, los modelos lineales en etapas y en particular los de dos etapas, permiten considerar incorporar información disponible de experimentos previos en la experimentación actual.

Los modelos en p -etapas fueron definidos por primera vez en 1986 por Lubomír Kubáček y a pesar de haber pasado más de 20 años desde su definición, aplicaciones específicas relacionadas con este tipo de modelos son difíciles de encontrar.

En este sentido, esta investigación pretende obtener expresiones de los estimadores del vector de medias y del vector de parámetros para el modelo en dos etapas regular, desde el punto de vista de los operadores de proyección y generalizar estos resultados a los modelos en p -etapas. Con este aporte se espera, además de facilitar el uso de este tipo de modelos para que pueda aprovecharse el gran potencial que poseen en muchas de las aplicaciones prácticas, destacar la importancia que los operadores de proyección tienen en la actualidad; particularmente en el área de los modelos lineales, al expresar al estimador del vector de medias de un modelo lineal, en la forma Py donde P es un proyector e y el vector de respuestas del modelo, por lo que los estimadores del vector de parámetros y de las funciones lineales paramétricas estimables se escriben como una función de la proyección del vector de respuesta y sobre el espacio de estimación del modelo lineal. Para el logro de esta meta, se pretende incorporar teorías relacionadas con las matrices de proyección ortogonal y los modelos lineales, desarrolladas por Rao (1974) y aquellas referidas a las funciones lineales cero, estudiadas por Bhimasankaran y Sengupta (1996), de modo que las expresiones de los estimadores obtenidos sean expresados como funciones de un operador de proyección oblicuo y el vector de respuestas y del modelo.

1.2. Objetivos

En esta investigación se pretende alcanzar los siguientes objetivos.

1.2.1. Objetivo General

Obtener expresiones alternativas de los estimadores del vector de parámetros y el vector de medias en el modelo lineal en 2-etapas regular, desde el punto de vista de los operadores de proyección y generalizar los resultados al modelo en p-etapas regular.

1.2.2. Objetivos Específicos

1. Encontrar estimadores para el vector de medias y el vector de parámetros del modelo lineal en 2-etapas regular, desde el enfoque de los operadores de proyección.
2. Generalizar los resultados del modelo en 2-etapas regular al modelo en p-etapas regular. Esto es, hallar expresiones de los estimadores para el vector de medias y el vector de parámetros del modelo lineal en p-etapas regular, desde el enfoque de los operadores de proyección.

1.3. Antecedentes e Importancia de la Investigación

Los modelos de regresión y de diseño están entre los modelos lineales que más han sido estudiados y su aplicación ha sido extensa, en diversas áreas. Su popularidad se debe fundamentalmente a la utilidad que han tenido, al explicar de manera sencilla, el comportamiento de una variable respuesta en función de una o más variables explicativas, a la gran cantidad de literatura disponible existente, de fácil comprensión para la mayoría de los interesados en el tema y a la facilidad con la que este tipo de modelos puede ajustarse a un conjunto de datos. Esto explica la diversidad de software disponible para analizarlos.

Entre los tipos de modelos lineales se encuentran los modelos en etapas. Según Kubáček (1986) un modelo de regresión $y = X\beta + \varepsilon$ es un modelo en p-etapas si tiene la siguiente estructura:

$$\left\{ \begin{array}{l} y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_p \end{bmatrix}, X = \begin{bmatrix} X_1 & 0 & 0 & \cdots & 0 \\ C_{2,1} & X_2 & 0 & \cdots & 0 \\ C_{3,1} & C_{3,2} & X_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ C_{p,1} & C_{p,2} & C_{p,3} & \cdots & X_p \end{bmatrix}, \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \\ \beta_p \end{bmatrix} \\ \Sigma = \begin{bmatrix} \Sigma_{1,1} & 0 & \cdots & 0 \\ 0 & \Sigma_{2,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & & \Sigma_{p,p} \end{bmatrix} \end{array} \right.$$

donde $y_1, y_2, \dots, y_p$ son vectores de observaciones de la variable respuesta, $\beta_1, \beta_2, \dots, \beta_p$ vectores de parámetros, $X_i, C_{i,j}, \Sigma_{i,j}$ matrices de dimensiones apropiadas, y si además se cumple la condición $\mathcal{M}(C'_{i,j}) \subset \mathcal{M}(X'_j)$, donde $\mathcal{M}(C'_{i,j})$ y $\mathcal{M}(X'_j)$ son los espacios vectoriales generados por las columnas de las matrices $C'_{i,j}$ y X'_j respectivamente, según la definición dada al comienzo del capítulo 2, $i = 1, 2, \dots, p, j = 1, 2, \dots, i-1$.

Esta definición es bastante general, al no especificarse supuestos sobre la distribución del vector de respuesta y , el rango de la matriz X , ni sobre la estructura de la matriz Σ .

El objetivo fundamental del trabajo de Kubáček (1986), fue encontrar las condiciones bajo las cuales los estimadores de los vectores de parámetros $\hat{\beta}_i(\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_{i-1}, y_i) = \hat{\beta}_i(y_1, y_2, \dots, y_i)$; esto es, las condiciones bajo las cuales el BLUE $\hat{\beta}_i(y_1, y_2, \dots, y_i)$ obtenido a partir de las observaciones de los vectores $y_1, y_2, \dots, y_i$ es igual al BLUE $\hat{\beta}_i(\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_{i-1}, y_i)$ basado en los estimadores $\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_{i-1}$, y el vector y_i , $i = 2, 3, \dots, p$.

Posterior al artículo de Kubáček, Volaufova (1987) estudia el modelo en 2-etapas bajo condiciones específicas. Concretamente, estudia el modelo:

$$y_1 = X_1 \beta_1 + \varepsilon_1$$

$$y_2 = X_2 \beta_2 + D \beta_1 + \varepsilon_2$$

con $E(\varepsilon_1) = 0$, $E(\varepsilon_2) = 0$, $E(\varepsilon_1 \varepsilon_1') = \sigma_1^2 H_1$, $E(\varepsilon_2 \varepsilon_2') = \sigma_2^2 H_2$, ε_1 , ε_2 no correlacionados y asume que los vectores y_1 , y_2 poseen distribución normal multivariante, las matrices D , X_1 , X_2 son conocidas, X_1 , X_2 de rango completo por columnas y H_1 , H_2 no singulares, el cual matricialmente se escribe como:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ D & X_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix} \quad (1.1)$$

Volaufova transforma la segunda etapa del modelo a partir del BLUE $\hat{\beta}_1 = Qy_1$, de la primera etapa, con $QX_1 = I$. La transformación que lleva a cabo es la siguiente:

$$\begin{aligned} y_2 &= X_2 \beta_2 + D \beta_1 + \varepsilon_2 \\ y_2 &= DQy_1 - DQy_1 + X_2 \beta_2 + D \beta_1 + \varepsilon_2 \\ y_2 - DQy_1 &= -DQy_1 + X_2 \beta_2 + D \beta_1 + \varepsilon_2 \\ y_2^* &= X_2 \beta_2 + D \beta_1 - DQy_1 + \varepsilon_2 \\ y_2^* &= X_2 \beta_2 + DQX_1 \beta_1 - DQy_1 + \varepsilon_2 \\ y_2^* &= X_2 \beta_2 - DQ(y_1 - X_1 \beta_1) + \varepsilon_2 \\ y_2^* &= X_2 \beta_2 - DQ\varepsilon_1 + \varepsilon_2 \end{aligned}$$

obteniéndose así el modelo transformado:

$$\begin{bmatrix} y_1 \\ y_2^* \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} I & 0 \\ -DQ & I \end{bmatrix} \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix} \quad (1.2)$$

con matriz de covarianza dada por:

$$\begin{aligned} Cov(y_1, y_2^*) &= V_\theta = \begin{bmatrix} I & 0 \\ -DQ & I \end{bmatrix} \begin{bmatrix} \sigma_1^2 H_1 & 0 \\ 0 & \sigma_2^2 H_2 \end{bmatrix} \begin{bmatrix} I & -Q'D' \\ 0 & I \end{bmatrix} \\ &= \sigma_1^2 \begin{bmatrix} H_1 & -C' \\ -C & CH_1^{-1}C' \end{bmatrix} + \sigma_2^2 \begin{bmatrix} 0 & 0 \\ 0 & H_2 \end{bmatrix}; C = DQH_1 \end{aligned}$$

Luego a partir de este modelo transformado, encuentra expresiones para los mejores

estimadores lineales insesgados localmente (LBLUE) de $\beta = (\beta'_1, \beta'_2)'$. Para ello utiliza la expresión

$$\tilde{\beta} = (\tilde{\beta}'_1, \tilde{\beta}'_2)' = (X'V_\theta^{-1}X)^{-1}X'V_\theta^{-1}y$$

siendo

$$X = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix}; y = \begin{bmatrix} y_1 \\ y_2^* \end{bmatrix}$$

resultado que establece en el Teorema 1.1. En este teorema, Volaufova expresa que los LBLUE $(\tilde{\beta}'_1, \tilde{\beta}'_2)'$ obtenidos a partir del modelo transformado son

$$\begin{aligned} \tilde{\beta}_1 &= V_\rho^{-1}V_\rho\hat{\beta}_1 + V_\rho^{-1}D'H_2^{-1}y_2^* + V_\rho^{-1}D'H_2^{-1}X_2E^{-1}X_2'H_2^{-1}(DV_\rho^{-1}D'H_2^{-1} - I)y_2^* \\ \tilde{\beta}_2 &= E^{-1}X_2'H_2^{-1}y_2^* - E^{-1}X_2'H_2^{-1}DV_\rho^{-1}D'H_2^{-1}y_2^* \end{aligned}$$

donde E , V_ρ y ρ están dados por

$$\begin{aligned} E &= X_2'H_2^{-1}X_2 - X_2'H_2^{-1}DV_\rho^{-1}D'H_2^{-1}X_2 \\ V_\rho &= \rho X_1'H_1^{-1}X_1 + D'H_2^{-1}D \\ \rho &= \sigma_2^2/\sigma_1^2 \end{aligned}$$

Posteriormente, en el Teorema 2.2 del mismo artículo, Volaufova demuestra que los mejores estimadores lineales insesgados uniformemente (UBLUE) del vector de parámetros $\beta = (\beta'_1, \beta'_2)'$ bajo el modelo transformado existen si y sólo si la condición $\mathcal{M}(D) \subset \mathcal{M}(X_2)$ se cumple, en cuyo caso los estimadores LBLUE también son UBLUE.

Esta condición fue mencionada por Kubáček (1986, p.90), y señala que la condición $\mathcal{M}(C_{i,j}) \subset \mathcal{M}(X_2)$ es importante en los modelos en dos etapas, haciendo referencia al artículo de Volaufova (1987) que fue publicado un año posterior al de él. En el Teorema 2.1, Volaufova (1987) también da expresiones para los mejores estimadores invariantes e insesgados de los parámetros σ_1^2, σ_2^2 .

A pesar de que Volaufova (1987) estudió al modelo de regresión en 2-etapas regular, fue Kubáček (1988) quien le da esta denominación en la Definición 1.1 de su artículo,

cuando expresa que el modelo de regresión en 2-etapas es regular si los rangos de las matrices X_1 , X_2 , Σ_{11} y Σ_{22} son k_1, k_2, n_1, n_2 respectivamente, donde X_1 es de dimensión $n_1 \times k_1$, X_2 de $n_2 \times k_2$, Σ_{11} de $n_1 \times n_1$ y Σ_{22} de $n_2 \times n_2$, implicando que las matrices X_1 , X_2 son de rango completo por columnas y Σ_{11} , Σ_{22} son no singulares.

Kubáček (1988) resalta que las matrices de covarianza son de la forma $\Sigma_{11} = \sigma_1^2 H_1$ y $\Sigma_{22} = \sigma_2^2 H_2$, donde $\sigma_i^2 \in (0, \infty)$, $i = 1, 2$ son los componentes de varianza. Los resultados principales los da en los teoremas 2.1 al 2.3 de su artículo. En el Teorema 2.1, asume que los componentes de varianza son conocidos y da varias expresiones para los estimadores LBLUE: $\hat{\beta}_1(y_1)$ que depende de y_1 , $\hat{\beta}_1(y_1, y_2)$ que depende de y_1, y_2 , y para $\hat{\beta}_2 \left[\hat{\beta}_1(y_1), y_2 \right]$ que depende de $\hat{\beta}_1(y_1)$ y de y_2 , y señala que en este caso $\hat{\beta}_2(y_1, y_2) = \hat{\beta}_2 \left[\hat{\beta}_1(y_1), y_2 \right]$.

En el Teorema 2.2, considera que los componentes de varianza σ_1^2, σ_2^2 y la razón $\rho = \sigma_2^2 / \sigma_1^2$ son desconocidos y especifica las condiciones que deben cumplirse para que los estimadores $\hat{\sigma}_1(y_1, y_2)$, $\hat{\sigma}_2(y_1, y_2)$ existan y sean LMVQUIE (estimadores insesgados invariantes cuadráticos de varianza mínima localmente), o sean UMVQUIE (estimadores insesgados invariantes cuadráticos de varianza mínima uniformemente). Luego en el Teorema 2.3, da expresiones para los estimadores de los componentes de varianza, sujetas a condiciones que deben cumplirse. Es de interés puntualizar, que para Kubáček (1988), UBLUE significa el BLUE uniforme con respecto a los parámetros de covarianza.

En este mismo año, Volaufova (1988) analiza el modelo en n-etapas

$$\begin{aligned} y_1 &= X_1 \beta_1 + \varepsilon_1 \\ y_2 &= C_{2,1} \beta_1 + X_2 \beta_2 + \varepsilon_2 \\ y_3 &= C_{3,1} \beta_1 + C_{3,2} \beta_2 + X_3 \beta_3 + \varepsilon_3 \\ &\vdots \\ y_n &= C_{n,1} \beta_1 + C_{n,2} \beta_2 + \cdots + C_{n,n-1} \beta_{n-1} + X_n \beta_n + \varepsilon_n \end{aligned}$$

donde para $i = 1, 2, \dots, n$, los y_i representan las mediciones de la variable respuesta, de $n_i \times 1$, en cada una de las i etapas del modelo, los cuales se asumen están distribuidos normalmente. Los β_i vectores de parámetros desconocidos a estimar, de dimensión

$k_i \times 1$. Las X_i matrices conocidas, de rango completo por columnas de $n_i \times k_i$, y las matrices $C_{i,j}$ también se suponen conocidas, de dimensión $n_i \times k_j$, y representan la conexión entre los parámetros de las diferentes etapas. En este modelo, los errores ε_i son no correlacionados, con $\varepsilon_i \sim N_{n_i}(0, \sigma_i^2 I_{n_i})$, y los componentes de varianza σ_i^2 , se consideran distintos, $i = 1, 2, \dots, n$.

En su trabajo, Volaufova (1988) enuncia varios resultados interesantes. En el Teorema 1 establece, que las condiciones necesarias y suficientes, que deben cumplirse para la existencia de los UMVUE de las funciones lineales estimables del vector de parámetros, para el modelo en n etapas anterior son:

$$\mathcal{M}(C_{i,j}) \subset \mathcal{M}(X_i) \forall j = 1, \dots, i-1; \forall i = 2, \dots, n$$

Para los demás teoremas de su artículo, Volaufova especifica que los $C_{i,j} = 0$ para $i = 3, \dots, n; j = 1, \dots, i-2$. En otras palabras, que sobre la etapa $i, i = 2, \dots$, únicamente existe influencia de la etapa $i-1$ anterior. En el Teorema 2 (p.44), da expresiones de los UMVUE para cada vector β_i del modelo, en función de las mediciones y_i de todas las etapas. Luego en el Teorema 2 (p.45), encuentra expresiones para los estimadores UMVUIE (estimadores insesgados invariantes de varianza mínima uniformemente), de cada componente de varianza σ_i^2 . Después en el Teorema 3, encuentra expresiones para los UMVUE de los vectores de parámetros $\beta_i, i = 1, 2, 3, 4$, para el caso particular de un modelo con cuatro etapas.

En 1993, Lubomír Kubáček analiza al modelo en dos etapas bajo restricciones y encuentra estimadores para el vector de parámetros β relacionado con la segunda etapa, cuando en el modelo se especifican restricciones de la forma $B\beta + C\theta + a = 0$, donde B, C son matrices conocidas, a un vector dado y θ el vector de parámetros de la primera etapa.

En 1996 Aiyi Liu estudia el sistema compuesto por dos modelos lineales con un vector de parámetros común β

$$\begin{aligned} y_1 &= X_1\beta + Z_1\beta_1 + \varepsilon_1 \\ y_2 &= X_2\beta + Z_2\beta_2 + \varepsilon_2 \end{aligned}$$

donde para $i = 1, 2$ y_i son vectores de observaciones, las matrices X_i, Z_i , conocidas de rango completo por columnas y los vectores de errores ε_i normales multivariantes con $E(\varepsilon_i) = 0, Cov(\varepsilon_i) = \sigma_i^2 I$

Liu encuentra estimadores BLUE para los vectores de parámetros, estudiando cada modelo en forma independiente y en forma conjunta, y en la Definición 1 de su trabajo señala que un estimador insesgado b_1 , de un parámetro β , es superior a otro estimador insesgado b_2 del mismo parámetro, si $Cov(b_1) \leq Cov(b_2)$; definición que utiliza para destacar que si σ_1^2, σ_2^2 son conocidas, el estimador del parámetro común β obtenido del sistema completo es superior a los obtenidos en forma independiente. Liu en su estudio también establece las condiciones en las cuales los BLUE de los vectores de parámetros obtenidos en forma conjunta, son superiores a los obtenidos en forma separada, en caso que las varianzas no sean conocidas.

El último trabajo que a la fecha se ha encontrado, relacionado específicamente con los modelos en etapas, fue realizado por Volaufova en 2004. En este artículo se dan las condiciones necesarias y suficientes para la existencia de estimadores UMVLUE (Estimadores Insesgados Lineales de Varianza Mínima Uniformemente), en el modelo de p-etapas con supuestos bastantes amplios. Como la misma autora expresa, "las condiciones necesarias y suficientes presentadas en el Teorema 1 no son muy transparentes y en la práctica puede que no sean fáciles de verificar" (p.393).

Otros autores han realizado estudios del modelo lineal desde otra óptica. Alalouf y Styán (1979) estudiaron el modelo imponiendo restricciones lineales de la forma $L'_0 \beta = 0$ y encontraron varios resultados interesantes relacionados con la estimación y prueba de hipótesis. Puntanen en 1996 y 1997 estudia el modelo lineal particionado $(y, X_1 \beta_1 + X_2 \beta_2, V)$ y su relación con los BLUE de algunos de sus sub modelos. Baksalary y Kala (1981), Drygas (1983) y Mueller (1987) estudian la suficiencia lineal en los modelos lineales.

En 2001 Gross et. al. establecen las condiciones que deben cumplirse para que $COLSE(X\beta) = CBLUE(X\beta)$ en el modelo de Gauss-Markov $\{y, X\beta, \sigma^2 V\}$ bajo tres circunstancias diferentes. Gross hace uso de los operadores de proyección ortogonales y los espacios vectoriales para establecer las condiciones en las cuales esta igualdad

ocurre. Luego en el 2004 GroB hace un recorrido por algunos resultados sobresalientes relacionados con el modelo de Gauss-Markov $\{y, X\beta, \sigma^2 V\}$ donde la matriz V puede ser deficiente en rango. En este artículo GroB asocia los vectores de respuestas y error a sus respectivos espacios vectoriales y da condiciones para la existencia de los BLUE de funciones estimables, los cuales establece son únicos, relacionándolos con proyectores oblicuos, entre otros resultados. Luego Zhang et. al. (2004) estudiaron el modelo lineal particionado y establecieron relaciones entre los BLUE de varios de sus sub modelos.

Más recientemente Christensen y Lin (2012) estudian el modelo $y = X\beta + \varepsilon$ para el caso cuando $\mathcal{M}(X) \not\subset \mathcal{M}(V)$, caso en el cual V es singular. En su trabajo ellos descomponen el vector de parámetros β como la suma de dos partes ortogonales $\beta = \beta_0 + \beta_1$, siendo β_0 conocido, y en lugar de analizar el modelo original, estudian el modelo $y - X\beta_0 = \tilde{X}\gamma + \varepsilon$ donde $\tilde{X}\gamma = X\beta_1$ la parte desconocida de $X\beta$, con $E(\varepsilon) = 0$, $Cov(\varepsilon) = \sigma^2 V$ y $\mathcal{M}(\tilde{X}) = \mathcal{M}(X) \cap \mathcal{M}(V)$, para el cual es claro que se cumple el supuesto $\mathcal{M}(\tilde{X}) \subset \mathcal{M}(V)$. Christensen y Lin en este trabajo muestran que este método proporciona los estimadores y pruebas de hipótesis usuales.

Los estudios sobre modelos lineales son muy diversos y constantemente aparecen nuevas publicaciones relacionadas con ellos, pero como ya se ha mencionado, investigaciones relacionados a los modelos en etapas, como los ha definido Kubáček no se han encontrado en las bases de datos al alcance de nuestras posibilidades. En este sentido, esta investigación pretende dar un aporte nuevo e inédito relacionado con este tipo de modelos, obteniendo expresiones para los estimadores del vector de medias y el vector de parámetros del modelo lineal en 2-etapas regular, desde el enfoque de los operadores de proyección, y luego generalizando los resultados obtenidos al modelo lineal en p-etapas regular. De esta manera, se espera renovar el interés sobre estos modelos al utilizar la metodología de los operadores de proyección y los espacios vectoriales, la cual ha despertado gran interés en el estudio de los modelos lineales, tal y como se evidencia en los artículos relacionados con los operadores de proyección, en particular el artículo de Christensen y Lin (2012). De esta forma, se espera que esta investigación sirva como punto de partida de futuras investigaciones.

Capítulo 2

Marco Teórico

2.1. Definiciones del álgebra lineal

Los espacios vectoriales y los operadores de proyección están íntimamente ligados al estimador del vector de medias $X\beta$ en un modelo lineal $y = X\beta + \varepsilon$, de aquí su importancia en la teoría de los modelos lineales.

Sea A una matriz $n \times p$ de números reales. Cada columna de A es un vector en R^n . Al espacio generado por las columnas de A se le denomina el espacio de columnas de A y se denota por $\mathcal{M}(A)$. Si B es una matriz $n \times r$, entonces $\mathcal{M}(A, B)$ es el espacio generado por las $p + r$ columnas de A y B . El rango de un espacio vectorial $\mathcal{M}(A)$ está dado por el número de columnas de A linealmente independientes y lo denotamos con $r(A)$.

Una definición importante en teoría de espacios vectoriales, es la de complemento ortogonal de un espacio vectorial, dada a continuación.

Definición 1 Para un subespacio N de M , sea $N_M^\perp = \{y \in M | y \perp N\}$. $N_M^\perp$ se denomina el complemento ortogonal de N con respecto a M . Si $M = R^n$, entonces a $N_M^\perp = N^\perp$ se conoce simplemente como el complemento ortogonal de N .

El complemento ortogonal de N con respecto a M es un subespacio de M y si $x \in M$, entonces puede escribirse como la suma $x = x_0 + x_1$ con $x_0 \in N$ y $x_1 \in N_M^\perp$. Los rangos de estos subespacios satisfacen la ecuación $r(M) = r(N) + r(N_M^\perp)$. En particular, si A

es una matriz cualquiera cuyas columnas pertenecen a R^n , al complemento ortogonal de $\mathcal{M}(A)$ se le denota con $\mathcal{M}(A)^\perp$. Además $\mathcal{M}(A)$ y el espacio nulo de A' son complementos ortogonales. El espacio nulo de A' se define como el conjunto de todos los x tales que $A'x = 0$ y se denota $\mathcal{N}(A')$. Si $r(A) = k$ se verifica que $r(\mathcal{N}(A')) = n - k$. En particular, si A es una matriz simétrica con $r(A) = k$ el complemento ortogonal de $\mathcal{M}(A)$ es $\mathcal{M}(A)^\perp = \mathcal{N}(A)$ y su rango es $r(\mathcal{N}(A)) = n - k$.

2.1.1. Operadores de proyección de un espacio vectorial

En la teoría de modelos lineales, los operadores de proyección desempeñan un rol fundamental. En general, el BLUE del vector $X\beta$ en el modelo de Gauss-Markov, puede expresarse como $\widehat{X\beta} = Py$, con P una matriz operador de proyección, de allí su importancia. En esta sección se tratarán tópicos relacionados con los operadores de proyección perpendiculares y oblicuos.

Definición 2 Sea X una matriz cualquiera y sea $\mathcal{M}(X)$ el espacio generado por las columnas de la matriz X . Una matriz A se dice es un operador de proyección perpendicular sobre $\mathcal{M}(X)$ si y sólo si se cumplen las siguientes dos condiciones

1. $v \in \mathcal{M}(X)$ implica que $Av = v$ (proyección)
2. $w \perp \mathcal{M}(X)$ implica que $Aw = 0$ (perpendicularidad)

Según Christensen (2011), el espacio de columnas generado por la matriz A es igual al espacio de columnas generado por la matriz X ; esto es, que $\mathcal{M}(X) = \mathcal{M}(A)$. Además, del Teorema B33 y la Proposición B34 de Christensen (2011, pág.427) se desprende que todo operador de proyección perpendicular posee tres propiedades fundamentales: unicidad, simetría e idempotencia. Únicos en el sentido que si M y P son operadores de proyección perpendiculares sobre el mismo espacio vectorial, entonces obligatoriamente se cumple que $M = P$. Simétricos e idempotentes ya que A es una matriz operador de proyección perpendicular sobre $\mathcal{M}(A)$ si y sólo si $A = A'$ (simetría) y $AA = A$ (idempotencia).

Dos maneras de caracterizar la matriz operador perpendicular sobre $\mathcal{M}(X)$, y algunos resultados relacionados con las matrices de proyección perpendicular son dados por Christensen (2011). Entre estos resultados se quieren destacar los siguientes:

Teorema 3 *Sea X una matriz cualquiera, entonces $X(X'X)^-X'$ es el operador de proyección perpendicular sobre $\mathcal{M}(X)$.*

A la matriz $X(X'X)^-X'$ la denotaremos con P_X para indicar que es el operador de proyección perpendicular sobre el espacio de columnas generado por la matriz X .

Teorema 4 *Si A y A_0 son matrices operador proyección perpendicular con $\mathcal{M}(A_0) \subset \mathcal{M}(A)$. Entonces $\mathcal{M}(A - A_0)$ es el complemento ortogonal de $\mathcal{M}(A_0)$ con respecto a $\mathcal{M}(A)$.*

La matriz $A - A_0$ es el operador proyección perpendicular sobre el complemento ortogonal de $\mathcal{M}(A_0)$ con respecto a $\mathcal{M}(A)$ y las matrices A_0 , A y $A - A_0$ satisfacen la relación $r(A) = r(A_0) + r(A - A_0)$. En el caso particular en que $A = I_n$ la matriz $I_n - A$ es el operador de proyección perpendicular sobre el complemento ortogonal de $\mathcal{M}(A)$ y se verifica que $n = r(A) + r(I_n - A)$.

Cualquier operador de proyección que no sea un operador de proyección perpendicular se conoce como operador de proyección oblicuo. A diferencia de los operadores de proyección perpendiculares, los oblicuos son matrices idempotentes solamente.

Definición 5 *Sean N y M dos espacios con $N \cap M = \{0\}$ y $r(M) + r(N) = n$. La matriz cuadrada A es un operador de proyección sobre N a lo largo de M si*

1. $Av = v$ para cualquier $v \in N$.
2. $Aw = 0$ para cualquier $w \in M$.

Dado que la matriz A satisface la ecuación $Av = v$ para cualquier $v \in N$, entonces $AAv = Av = v$, luego sólo es posible que $AA = A$ y en consecuencia A es idempotente. Si A es una matriz cuadrada con la propiedad $Av = v$ para cualquier $v \in \mathcal{M}(A)$, entonces a la matriz A se le conoce como el operador proyección sobre $\mathcal{M}(A)$ a lo largo de $\mathcal{M}(A')^\perp = \mathcal{N}(A) = \text{espacio nulo de } A$.

2.2. Los Modelos Lineales

En todo estudio estadístico con miras a realizar el ajuste de un modelo, siempre hay que seleccionar entre la sencillez y complejidad. Los modelos más sencillos tienden a ser más fáciles de entender y computacionalmente más manejables, pero el ajuste de los datos puede resultar poco satisfactorio. Por otro lado, los modelos más complejos tienden a ajustar mejor los datos, pero en la medida que su complejidad aumenta, son más difíciles de interpretar y manipular, además de ser difíciles de replicar. Por ello los procedimientos clásicos como el modelo lineal general han tenido un gran éxito, ya que para la mayoría de las aplicaciones en que se emplean, el ajuste de los datos resulta satisfactorio y computacionalmente manejable, de ahí su empleo en muchas aplicaciones científicas.

A pesar de la diversidad de herramientas computacionales existentes en la actualidad, que han permitido realizar análisis complejos como algoritmos iterativos, métodos de Monte Carlo, pruebas repetidas, estudios multivariantes, que ahora se realizan de manera rutinaria, los métodos clásicos siguen estando entre los mejores para dar respuesta a muchas interrogantes, además de ser la base de muchos de los desarrollos más recientes.

En estadística, un modelo lineal general se escribe como

$$Y = \mu(X) + \varepsilon$$

donde Y y ε son variables aleatorias, y a $\mu(X)$ se le conoce como la parte determinística del modelo. En general, este tipo de modelos se utilizan para determinar Y a partir del conocimiento de la variable X , definida en un conjunto D y que en este trabajo se considerará no aleatoria. A la variable Y se le conoce como la variable dependiente o variable respuesta y a X como la variable independiente o predictiva. Con frecuencia la forma funcional de $\mu(X)$ es conocida pero depende de parámetros desconocidos. El modelo se denomina lineal en el sentido que $\mu(X)$ es una función lineal de los parámetros desconocidos. Más específicamente, Graybill (1976) da la siguiente definición de modelo lineal general:

Definición 6 *El Modelo Lineal General.* Sean Y de $n \times 1$ un vector observable

de variables aleatorias; X una matriz de $n \times p$ ($n > p$) de números fijos conocidos; β un vector $p \times 1$ de parámetros desconocidos, y ε un vector no observable de variables aleatorias de $n \times 1$, donde $E(\varepsilon) = 0$ y $Cov(\varepsilon) = \Sigma$. Estas especificaciones relacionadas a través de la ecuación

$$Y = X\beta + \varepsilon \quad (2.1)$$

definen un modelo lineal general.

El caso más sencillo del modelo lineal general ocurre cuando $\Sigma = \sigma^2 I$, $\sigma^2 > 0$ y la matriz X es de rango completo por columnas.

2.3. Estimación

Las diferentes variantes del modelo (2.1) están relacionadas con la especificación de las propiedades distribucionales de ε o de Y , con la especificación de los supuestos acerca de la estructura de la matriz Σ , o con el rango de la matriz X .

2.3.1. Estimación por Máxima Verosimilitud

Supongamos que $\Sigma = \sigma^2 V$, con V una matriz conocida definida positiva y que ε tiene distribución normal multivariante con $E(\varepsilon) = 0$ y $Cov(\varepsilon) = \Sigma$. Entonces Y está distribuido $\mathcal{N}(y; X\beta, \sigma^2 V)$ y su función de verosimilitud está dada por

$$\begin{aligned} L(\beta, \sigma^2; y_1, \dots, y_n, x_{11}, \dots, x_{np}) &= L(\beta, \sigma^2; Y, X) \\ &= (2\pi\sigma^2)^{-n/2} |V|^{-1/2} \\ &\quad \exp \left[-\frac{1}{2\sigma^2} (Y - X\beta)' V^{-1} (Y - X\beta) \right] \end{aligned}$$

donde el espacio paramétrico está dado por $\Theta = \{(\beta, \sigma^2); \sigma^2 > 0, -\infty < \beta_i < \infty, i = 1, \dots, p\}$. El logaritmo de la verosimilitud es entonces

$$\begin{aligned} \log L(\beta, \sigma^2; Y, X) &= -\frac{n}{2} \log 2\pi - \frac{n}{2} \log \sigma^2 - \frac{1}{2} \log |V| - \frac{1}{2\sigma^2} (Y - X\beta)' V^{-1} (Y - X\beta) \\ &= -\frac{n}{2} \log 2\pi - \frac{n}{2} \log \sigma^2 - \frac{1}{2} \log |V| \\ &\quad - \frac{1}{2\sigma^2} (Y' V^{-1} Y - 2\beta' X' V^{-1} Y + \beta' X' V^{-1} X \beta) \end{aligned}$$

Para encontrar los estimadores de máxima verosimilitud de los parámetros en Θ , se igualan a cero las derivadas parciales del logaritmo de la verosimilitud,

$$\begin{aligned} \frac{\partial \log L(\beta, \sigma^2; Y, X)}{\partial \beta} &= \frac{1}{\sigma^2} (X' V^{-1} Y + X' V^{-1} X \beta) = 0 \\ \frac{\partial \log L(\beta, \sigma^2; Y, X)}{\partial \sigma^2} &= -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} (Y - X\beta)' V^{-1} (Y - X\beta) = 0 \end{aligned}$$

obteniéndose las conocidas ecuaciones normales, dadas por

$$\begin{aligned} X' V^{-1} X \tilde{\beta} &= X' V^{-1} Y \\ \tilde{\sigma}^2 &= \frac{1}{n} (Y - X \tilde{\beta})' V^{-1} (Y - X \tilde{\beta}) \end{aligned} \quad (2.2)$$

donde $\tilde{\beta}$ y $\tilde{\sigma}^2$ son los estimadores de máxima verosimilitud de β y σ^2 respectivamente.

De acuerdo al Teorema 14.11.1 de Harville (1997, p.256), el sistema de ecuaciones $X' V^{-1} X \tilde{\beta} = X' V^{-1} Y$ es consistente en β . Luego, del Teorema 9.1.2 de Harville (1997, p.108), una solución para $\tilde{\beta}$ está dada por $\tilde{\beta} = (X' V^{-1} X)^{-} X' V^{-1} Y$. Si el rango de la matrix X es p ; esto es, si $r(X) = p$, se sigue que $r(X' V^{-1} X) = p$, en consecuencia $(X' V^{-1} X)^{-1}$ existe, y en este caso la única solución es $\tilde{\beta} = (X' V^{-1} X)^{-1} X' V^{-1} Y$.

Cuando la matrix X no es de rango completo por columnas existen infinitas soluciones para el vector de parámetros β . En este caso el valor esperado de $\tilde{\beta}$ está dado por

$$E(\tilde{\beta}) = (X' V^{-1} X)^{-} X' V^{-1} X \beta$$

Cuando $r(X) = p$, el resultado que se obtiene para el valor esperado es

$$E(\tilde{\beta}) = (X'V^{-1}X)^{-1}X'V^{-1}X\beta = \beta$$

y sólo en este caso $\tilde{\beta} = (X'V^{-1}X)^{-1}X'V^{-1}Y$ es un estimador insesgado para el vector de parámetros β , el cual se denotará con $\hat{\beta}$, esto es,

$$\hat{\beta} = (X'V^{-1}X)^{-1}X'V^{-1}Y \quad (2.3)$$

En lo que respecta al parámetro σ^2 , al sustituir $\tilde{\beta}$ en la expresión (2.2) se obtiene:

$$\hat{\sigma}^2 = \frac{1}{n}Y' [V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}] Y$$

Como la matriz V es definida positiva, se puede descomponer como el producto $V = PP'$ con P una matriz no singular (Christensen, 2011, p.424). De este modo, se escribe

$$\begin{aligned} \hat{\sigma}^2 &= \frac{1}{n}Y' [(PP')^{-1} - (PP')^{-1}X(X'(PP')^{-1}X)^{-1}X'(PP')^{-1}] Y \\ &= \frac{1}{n}Y' P'^{-1} [I - P^{-1}X(X'P'^{-1}P^{-1}X)^{-1}X'P'^{-1}] P^{-1}Y \end{aligned}$$

y al denotar $A = P^{-1}X$, y $Z = P^{-1}Y$ se obtiene

$$\hat{\sigma}^2 = \frac{1}{n}Z' [I - A(A'A)^{-1}A'] Z$$

Obsérvese que como $E(Z) = P^{-1}X\beta = A\beta$ y $Cov(Z) = \sigma^2 P^{-1}PP'P'^{-1} = \sigma^2 I$ se sigue, del Teorema 1.3.2 de Christensen (2011, p.8), del Lema 10.2.4 de Harville (1977, p.135) y al hecho que $A(A'A)^{-1}A'$ es el operador proyección sobre el espacio columna

generado por la matriz A , que

$$\begin{aligned} E(Z' [I - A(A'A)^{-1}A'] Z) &= \text{tr}([I - A(A'A)^{-1}A'] \sigma^2 I) + \beta' A' [I - A(A'A)^{-1}A'] A \beta \\ &= \sigma^2 r([I - A(A'A)^{-1}A']) = \sigma^2(n - r[A(A'A)^{-1}A']) \\ &= \sigma^2(n - r(A)) = \sigma^2(n - r(X)) \end{aligned}$$

y entonces, es sencillo ver que un estimador insesgado para σ^2 está dado por $\hat{\sigma}^2 = n\tilde{\sigma}^2/(n - r(X))$. Específicamente

$$\begin{aligned} \hat{\sigma}^2 &= \frac{1}{n - r(X)} (Y - X\tilde{\beta})' V^{-1} (Y - X\tilde{\beta}) \\ \hat{\sigma}^2 &= \frac{1}{n - r(X)} Y' [V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}] Y \end{aligned} \quad (2.4)$$

independientemente del valor particular de $\tilde{\beta}$ que se utilice. Cuando $r(X) = p$, la expresión para $\hat{\sigma}^2$ es prácticamente la misma, sólo que en este caso, la inversa de $X'V^{-1}X$ existe y por tanto

$$\hat{\sigma}^2 = \frac{1}{n - p} Y' [V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}] Y \quad (2.5)$$

2.3.2. Estimadores UMVU para β y σ^2

Cuando se desea estimar un parámetro, siempre es deseable que el estimador cumpla con ciertas características que lo hagan el mejor, entre todos los posibles estimadores. En estadística, de ser posible, se utilizan estimadores insesgados de varianza mínima uniformemente (UMVUE) para realizar la estimación, ya que cuando existen, son los mejores. Graybill (1976, págs. 73,78-79) define de manera clara los UMVUE y da un teorema para encontrarlos en caso de que se disponga de un conjunto de estadísticos suficientes y completos para el parámetro a estimar (el teorema de Lehmann-Scheffé). Resultados que se enuncian a continuación.

Definición 7 *Estimador Insesgado de Varianza Mínima Uniformemente (UMVUE).* Sea $Y_1, Y_2, \dots, Y_n$ una muestra observable con f.d.p. $f_Y(y : \theta)$, $\theta \in \Theta$. Sea

$t(Y)$ un estimador de $c(\theta)$; entonces $t(Y)$ es un estimador insesgado de varianza mínima uniformemente (UMVUE) para $c(\theta)$ si y sólo si $t(Y)$ satisface (1) y (2) a continuación:

1. $E[t(Y)] = c(\theta)$, esto es, $t(Y)$ es un estimador insesgado de $c(\theta)$.
2. $Var[t(Y)] \leq Var[t^*(Y)]$, donde $t^*(Y)$ es cualquier otro estimador insesgado de $c(\theta)$.

Teorema 8 Teorema de Lehmann-Scheffé. Sea $Y_1, Y_2, \dots, Y_n$ una muestra observable con f.d.p. $f_Y(y : \theta)$, $\theta \in \Theta$. Sean $S_1 = s_1(Y_1, Y_2, \dots, Y_n), \dots, S_k = s_k(Y_1, Y_2, \dots, Y_n)$ un conjunto de estadísticos suficientes y completos. Sea $c(\theta)$ cualquier función conocida de θ , y supóngase que $T = t(S_1, S_2, \dots, S_n)$ es un estimador insesgado de $c(\theta)$. Entonces T es el (único) estimador UMVU de $c(\theta)$.

Graybill (1976, p. 208) también expresa que los estimadores (2.3) y (2.5) son estadísticos suficientes y completos para β y σ^2 , en consecuencia, del Teorema 8 se sigue que ellos son UMVUE. El caso particular cuando $V = I_n$ es ampliamente conocido. Para más detalle puede consultarse el Teorema 6.2.1 de Graybill(1976).

2.3.3. Estimadores de Mínimos Cuadrados

Según Christensen (2011, p.17), el método de estimación de mínimos cuadrados se remonta al menos a 1805, la idea se le atribuye a Legendre, pero Gauss afirmó haber inventado el método antes de 1805. En todo caso, el método de mínimos cuadrados ha resultado ser una herramienta muy utilizada en teoría de la estimación. Este método brinda una gran alternativa para encontrar estimadores de parámetros en aquellos casos en los que el método de máxima verosimilitud no se puede aplicar, por desconocerse la distribución de probabilidad de la población.

Considérese de nuevo el modelo lineal general (2.1) con $\Sigma = \sigma^2 I$. Como no se conoce nada acerca de la distribución del vector Y , aplicamos el método de mínimos cuadrados. En este caso, el método produce como resultado al vector β que minimiza la suma de los cuadrados de las diferencias entre el verdadero valor de una observación y_i y el valor

estimado por el modelo $\hat{y}_i$. Esto es, el objetivo del método es encontrar el estimador de mínimos cuadrados $\hat{\beta}_{LS}$ para el cual la siguiente ecuación se cumple

$$\hat{\beta}_{LS} = \min_{\beta} \sum_{i=1}^n e_i^2 = \min_{\beta} (Y - X\beta)'(Y - X\beta)$$

donde $e_i = y_i - \hat{y}_i$. En palabras, $\hat{\beta}_{LS}$ es el valor del vector de parámetros β que minimiza $\sum_{i=1}^n e_i^2 = (Y - X\beta)'(Y - X\beta)$. Por ello, para obtenerlo se debe diferenciar esta suma de cuadrados con respecto a β y luego igualar las derivadas a cero. Al hacer esto, se obtiene

$$\frac{\partial \sum_{i=1}^n e_i^2}{\partial \beta} = \frac{\partial (Y - X\beta)'(Y - X\beta)}{\partial \beta} = -2X'Y + 2X'X\beta = 0$$

luego

$$X'X\beta = X'Y \quad (2.6)$$

A la ecuación (2.6) se le denomina la ecuación normal y una solución para esta ecuación es:

$$\tilde{\beta}_{LS} = (X'X)^- X'Y \quad (2.7)$$

donde $(X'X)^-$ es cualquier g-inversa de $X'X$. Si la matriz X es de rango completo por columnas, $(X'X)^{-1}$ existe y en este caso particular, el único estimador de mínimos cuadrados está dado por:

$$\hat{\beta}_{LS} = (X'X)^{-1} X'Y$$

La suma de cuadrados del error no depende de σ^2 , y por ello este método no proporciona un estimador para este parámetro. En este caso, lo que se acostumbra es utilizar como estimador para σ^2 a

$$\hat{\sigma}_{LS}^2 = \frac{1}{n - r(X)} (Y - X\hat{\beta}_{LS})' (Y - X\hat{\beta}_{LS}) \quad (2.8)$$

el cual se obtiene al sustituir el estimador $\hat{\beta}_{LS}$ en la expresión (2.4), con $V = I$ en este caso. Similar a la sección 2.3.1, el estimador (2.8) es invariante de la solución de mínimos cuadrados $\hat{\beta}_{LS}$ que se utilice. Obsérvese que si $Y \sim \mathcal{N}(y; X\beta, \sigma^2 I)$, el estimador dado

en (2.8) es igual al estimador insesgado $\hat{\sigma}^2$, y si además $r(X) = p$, el estimador obtenido en (2.7) también coincide con su par de máxima verosimilitud.

2.4. Identificabilidad, estimabilidad y estimadores BLUE

Además de los parámetros del modelo, en un estudio existen funciones lineales de los parámetros cuya estimación puede ser de interés en la investigación. En este sentido es necesario conocer cuando una función paramétrica se puede estimar, sobre todo cuando los parámetros del modelo carecen de una solución única. Graybill (1976, pág.217-218) señala que no es posible hacer aseveraciones fuertes relacionadas con las propiedades de los estimadores de mínimos cuadrados, a causa del desconocimiento de la distribución de ε (y de Y) y que en esta situación, lo que puede hacerse es exigir que los estimadores cumplan con la propiedad de varianza mínima y limitarlos a que pertenezcan a la clase de las funciones lineales, para luego determinar si dentro de esta clase restringida existen estimadores lineales insesgados de varianza mínima uniformemente (UMVBLUE), también conocidos bajo el nombre de BLUE por ser los mejores (en el sentido que tienen varianza mínima) estimadores lineales insesgados (ver nota 1 de Graybill (1976, p.218).

En este contexto es necesario comprender cuando un estimador es BLUE, así como tener un conocimiento claro de dos conceptos fundamentales en la teoría de los modelos lineales, como lo son la identificabilidad y la estimabilidad. Las funciones lineales de los parámetros que son identificables, son las que pueden estimarse y sus estimadores son lineales e insesgados, por ello reciben el nombre de funciones estimables. De ahí la importancia de estos conceptos.

Considérese ahora el modelo

$$Y = X\beta + \varepsilon \quad (2.9)$$

donde Y es un vector de observaciones de $n \times 1$, X una matriz $n \times p$ de constantes conocidas, β un vector de parámetros desconocidos de $p \times 1$ y ε un vector de errores aleatorio no observable de $n \times 1$ cuya distribución no depende de β .

Se denomina una parametrización del vector de medias $E(Y)$ a una función $f(\beta)$ del vector de parámetros β del modelo. Se escribe $E(Y) = f(\beta)$ para indicar que $f(\beta)$ es la parametrización $E(Y)$. En el caso del modelo lineal la parametrización de $E(Y) = X\beta$ es una función lineal del parámetro β .

Definición 9 Identificabilidad. *El parámetro β es identificable si para cualquier β_1 y β_2 , $f(\beta_1) = f(\beta_2)$ implica que $\beta_1 = \beta_2$. Si β es identificable, decimos que la parametrización $f(\beta)$ es identificable. Además, una función vector -valuada $g(\beta)$ es identificable si $f(\beta_1) = f(\beta_2)$ implica que $g(\beta_1) = g(\beta_2)$. Si la parametrización no es identificable pero existen funciones identificables no triviales, entonces la parametrización se dice es parcialmente identificable.*

En los modelos de regresión, la matriz X es de rango completo por columnas, por lo que $X'X$ es no singular, en consecuencia si $X\beta_1 = X\beta_2$, resulta que $\beta_1 = (X'X)^{-1}X'X\beta_1 = (X'X)^{-1}X'X\beta_2 = \beta_2$ y entonces el parámetro β y la parametrización $E(Y) = X\beta$ son identificables. Si $r(X) < p$ existen vectores $\beta_1 \neq \beta_2$ para los cuales $X\beta_1 = X\beta_2$ y entonces se dice que el vector β es no identificable, sin embargo $X\beta$ es identificable.

Con frecuencia se necesita saber si una función del parámetro β es identificable, ya que como se señaló anteriormente, si es identificable es estimable. En este sentido, el siguiente resultado permite comprender de forma bastante clara cuando una función $g(\beta)$, del parámetro β , es identificable.

Teorema 10 *Una función $g(\beta)$ es identificable si y sólo si $g(\beta)$ es una función de $f(\beta)$.*

De esta manera, en los modelos lineales la parametrización $f(\beta) = X\beta$ es la que determina cuando una función paramétrica $g(\beta)$ es o no identificable. De acuerdo a este teorema $g(\beta) = f(\beta) = X\beta$ siempre es identificable, lo cual tiene completo sentido ya que en el modelo lineal $X\beta$ representa a $E(Y)$. Por ello en los modelos lineales toda $g(\beta)$ que sea función de $X\beta$ es identificable. Cuando $r(X) = p$ el vector β podemos escribirlo $g(\beta) = h[f(\beta)] = (X'X)^{-1}X'X\beta = (X'X)^{-1}X'f(\beta)$ y por lo tanto β es identificable. Christensen (2011) refiere que no es razonable estimar funciones que sean no identificables, debido a que en estos casos no se sabe qué es lo que se estima.

La definición a continuación es sin lugar a dudas una de la más importantes en la teoría de la estimación en los modelos lineales, ya que caracteriza a las funciones lineales del parámetro β que pueden ser estimables.

Definición 11 Una función lineal del vector β , denotada $\Lambda'\beta$, es estimable si $\Lambda'\beta = P'X\beta$ para alguna matriz P .

La condición $\Lambda'\beta = P'X\beta$ implícitamente expresa que $\Lambda'\beta$ es identificable. Si $P' = I$, es evidente que $f(\beta) = E(Y) = X\beta$ siempre es estimable. Si $P' = l'_i$, donde l_i es un vector de $p \times 1$ con todos sus elementos iguales a cero, excepto el elemento de la posición i para $i = 1, \dots, p$, entonces $P'X\beta = l'_iX\beta$ que representa el producto de la fila i de la matriz X con el vector β , es estimable. Coloquialmente se dice que "las filas de la matriz X siempre son estimables".

Si $\Lambda'\beta = P'X\beta$ entonces $\Lambda' = P'X$ o equivalentemente $\Lambda = X'P$ y en consecuencia Λ pertenece al espacio generado por las columnas de X' ; esto es, $\Lambda \in \mathcal{M}(X')$. De esta manera, una función lineal paramétrica $\Lambda'\beta$ es estimable si $\Lambda \in \mathcal{M}(X')$.

Si $\Lambda'\beta$ es una función lineal estimable se cumple $\Lambda'\beta = \rho'X\beta$, pero como $E(Y) = X\beta$, escribimos $\Lambda'\beta = \rho'E(Y) = E(\rho'Y)$. Por otra parte, si $E(\rho'Y) = \Lambda'\beta$ se sigue que $\rho'X\beta = \Lambda'\beta$ y en este caso $\Lambda'\beta$ sería una función de $X\beta$ y por lo tanto una función estimable. Estas relaciones se mantienen para cualquier β , y se establecen en el siguiente resultado.

Proposición 12 $\Lambda'\beta$ es una función lineal estimable si y sólo si $E(\rho'Y) = \Lambda'\beta$ para cualquier β .

Christensen (2011) en la sección 2.2 da resultados muy importantes relacionados con las funciones estimables y el estimador de mínimos cuadrados $\hat{\beta}_{LS}$. Entre estos establece que $X\hat{\beta}_{LS}$ es el vector en $\mathcal{M}(X)$ más cercano al vector Y , al ser $X\hat{\beta}_{LS}$ la proyección perpendicular del vector Y sobre $\mathcal{M}(X)$. También demuestra que el estimador de mínimos cuadrados de la función lineal paramétrica $\rho'X\beta$ es $\rho'X\hat{\beta}_{LS}$, es único e insesgado. Adicionalmente, prueba que el método de mínimos cuadrados proporciona estimadores únicos de funciones identificables y estimadores insesgados de funciones estimables, indiferentemente de la estructura de covarianza.

Los conceptos de identificabilidad y de funciones lineales estimables, permiten tener una visión clara de la relación existente entre los estimadores de las funciones lineales estimables y el estimador de mínimos cuadrados para β . En consecuencia, es posible enunciar dos resultados que contribuirán a ampliar el conocimiento de esta relación. Para hacer esto es necesario agregarle al modelo (2.9) la condición $Cov(\varepsilon) = \sigma^2 I$. Específicamente estos resultados son la definición de mejor estimador lineal insesgado (BLUE) y el Teorema de Gauss-Markov. Este último es el que establece de manera contundente la importancia de los estimadores de mínimos cuadrados en la teoría de estimación de las funciones lineales paramétricas. Ambos pueden consultarse en Christensen (2011, p.28).

Definición 13 *Mejor Estimador Lineal Insesgado* $a'Y$ es un mejor estimador lineal insesgado de $X'\beta$ si $a'Y$ es insesgado y si para cualquier otro estimador lineal insesgado $b'Y$, $Var(a'Y) \leq Var(b'Y)$.

Teorema 14 *Teorema de Gauss-Markov.* Considérese el modelo lineal $Y = X\beta + \varepsilon$, donde $E(\varepsilon) = 0$, $Cov(\varepsilon) = \sigma^2 I$. Si $X'\beta$ es estimable, entonces el estimador de mínimos cuadrados de $X'\beta$ es un BLUE de $X'\beta$.

2.4.1. Estimadores de Mínimos Cuadrados Generalizados

Considérese de nuevo el modelo (2.1) con $\Sigma = \sigma^2 V$ y V definida positiva. Sea

$$Y = X\beta + \varepsilon \quad (2.10)$$

con $E(\varepsilon) = 0$ y $Cov(\varepsilon) = \Sigma$. De acuerdo al corolario B.23 de Christensen (2011) se puede escribir $V = QQ'$ para alguna matriz Q no singular y en consecuencia transformar al modelo $Y = X\beta + \varepsilon$ en el modelo

$$\begin{aligned} Q^{-1}Y &= Q^{-1}X\beta + Q^{-1}\varepsilon \\ Y^* &= X^*\beta + \varepsilon^* \end{aligned} \quad (2.11)$$

para el cual se cumple

$$\begin{aligned} E(\varepsilon^*) &= E(Q^{-1}\varepsilon) = 0 \\ \text{Cov}(\varepsilon^*) &= \text{Cov}(Q^{-1}\varepsilon) = Q^{-1}\sigma^2 V Q^{-1'} = \sigma^2 Q^{-1} Q Q' Q^{-1} = \sigma^2 I \end{aligned}$$

y al aplicar los resultados de la sección 2.3.3 se obtiene al estimador

$$\begin{aligned} \hat{\beta}_{GLS} &= (X^{*'} X^*)^{-1} X^{*'} Y^* = (X' Q'^{-1} Q^{-1} X)^{-1} X' Q'^{-1} Q^{-1} Y \\ &= (X' (Q Q')^{-1} X)^{-1} X' (Q Q')^{-1} Y \\ &= (X' V^{-1} X)^{-1} X' V^{-1} Y \end{aligned} \tag{2.12}$$

al cual se le denomina estimador de mínimos cuadrados generalizados, para no confundirlo con el dado en la ecuación (2.7), que se conoce bajo el nombre de estimador de mínimos cuadrados ordinarios. El estimador $\hat{\beta}_{GLS}$ está unívocamente definido si y sólo si X tiene rango completo por columnas.

Para finalizar esta sección, se expone un teorema dado por Christensen (2011), que relaciona a los estimadores de mínimos cuadrados de funciones lineales estimables con los UMVUE.

Teorema 15

- (a) $X'\beta$ es estimable en el modelo (2.10) si y sólo si $X'\beta$ es estimable en el modelo (2.11).
- (b) $\hat{\beta}$ es un estimador de mínimos cuadrados generalizados de $X\beta$ si y sólo si

$$X(X'V^{-1}X)^{-1}X'V^{-1}Y = X\hat{\beta}.$$

Para cualquier función estimable existe un único estimador de mínimos cuadrados generalizados.

- (c) Para una función estimable $X'\beta$, el estimador de mínimos cuadrados generalizados es el BLUE de $X'\beta$.

- (d) Si $\varepsilon \sim \mathcal{N}(0, \sigma^2 V)$, entonces para cualquier función estimable $\lambda' \beta$, el estimador de mínimos cuadrados generalizados es el estimador insesgado de varianza mínima UMVU.
- (e) Si $\varepsilon \sim \mathcal{N}(0, \sigma^2 V)$, entonces cualquier estimador de mínimos cuadrados generalizados de β es un estimador de máxima verosimilitud de β .

Es importante señalar que los items (b), (c), (d), y (e) son válidos cuando $Cov(\varepsilon) = \sigma^2 I$, sustituyendo la matriz V por la matriz identidad I , caso en el cual los estimadores de mínimos cuadrados generalizados y los estimadores de mínimos cuadrados ordinarios son iguales.

2.5. Estimadores UBLUE

En este trabajo se usa la definición de mejor estimador lineal insesgado uniformemente conocido como UBLUE, dada por Wulff y Birkes (2005).

Definición 16 *Considérese el modelo lineal $y = X\beta + \varepsilon$ con $Cov(y) = V_\psi$, donde X es una matriz conocida de $n \times p$, β un vector de parámetros desconocidos de efectos fijos de $p \times 1$, V_ψ una matriz definida positiva con ψ un vector de parámetros de covarianza desconocidos de $r \times 1$ los cuales pertenecen a un conjunto conocido Ψ . Un estimador de una función lineal paramétrica estimable se dice es UBLUE de su valor esperado si este es BLUE para todos los posibles valores de $\psi \in \Psi$.*

Según Wulff y Birkes (2005) bajo el supuesto de normalidad de los datos, los UBLUE son también estimadores insesgados de varianza mínima uniformemente conocidos como UMVUE.

2.6. Los operadores de proyección y los BLUE

La importancia de los operadores de proyección en la estimación del vector de medias del modelo general de Gauss-Markov $(y, X\beta, \sigma^2 V)$, fue puesta en evidencia por Rao (1974), al descomponer el vector de respuestas y como la suma $y = y_1 + y_2$ con $y_1 \in$

$\mathcal{M}(X)$, $y_2 \in \mathcal{M}(VZ)$, donde el vector y_1 es tanto el BLUE de $X\beta$ como la proyección de y sobre el subespacio $\mathcal{M}(X)$ a lo largo del subespacio $\mathcal{M}(VZ)$, y $Z = X^\perp$ una matriz de máximo rango tal que $X'Z = X'X^\perp = 0$, y $y_1 = Py$, donde P es un operador de proyección.

Dada la importancia que el artículo de Rao (1974) tiene en el desarrollo de este trabajo, se considera de interés destacar algunos de sus resultados relacionados con los operadores de proyección y sus representaciones. Adicionalmente, se dará la demostración realizada por el autor de este trabajo, exhibiéndose algunos pasos faltantes no dados en la demostración del artículo original.

Dos matrices A, B se dice son disjuntas si los espacios $\mathcal{M}(A)$ y $\mathcal{M}(B)$ son disjuntos; esto es, si $\mathcal{M}(A) \cap \mathcal{M}(B) = \{0\}$, resultado que se emplea en la demostración del lema siguiente.

Lema 17 Sea X una matriz $n \times m$, V una matriz n.n.d. de orden n , y sea $Z = X^\perp$. Entonces:

(i) X y VZ son matrices disjuntas.

(ii) $\mathcal{M}(V : X) = \mathcal{M}(VZ : X)$.

Demostración. Supóngase que $VZb = Xa$ para algunos vectores a y b . Entonces, al pre multiplicar ambos lados de la ecuación anterior por Z' , se obtiene

$$Z'VZb = Z'Xa = 0$$

ya que $Z'X = 0$. Luego $b'Z'VZb = (Zb)'V(Zb) = 0$ y por el Corolario 14.3.11 del Harville (1997) se sigue que $VZb = 0$. En consecuencia, la ecuación $VZb = Xa$ se cumple si $VZb = Xa = 0$. En otras palabras, $\mathcal{M}(VZ) \cap \mathcal{M}(X) = \{0\}$, lo cual demuestra que las matrices VZ y X son disjuntas.

De (i) se sigue que $\mathcal{M}(X)$ y $\mathcal{M}(VZ)$ son disjuntos, entonces

$$\begin{aligned}\mathcal{M}(VZ : X) &= \mathcal{M}(VZ) \oplus \mathcal{M}(X) \\ r(VZ : X) &= r(VZ) + r(X)\end{aligned}$$

Del Lema 12.5.2 y el Corolario 11.2.2 de Harville(1997), es conocido que $\mathcal{M}(X)^\perp = \mathcal{N}(X') = \mathcal{M}(I - (X')^-X')$. Entonces denominando $Z = I - (X')^-X'$, se cumple:

- $X'Z = X'(I - (X')^-X') = X' - X'(X')^-X' = X' - X' = 0.$
- $r(Z) = r(I - (X')^-X') = n - r(X') = n - r(X).$
- $r(X) + r(Z) = n.$

De lo anterior, es claro que $Z = I - (X')^-X'$ es una matriz de rango máximo que satisface $X'Z = 0$. Además, vemos que también se cumple $r(VZ) = r(Z'V) = r[(I - XX^-)V]$. Por otra parte, del Corolario 17.2.9 con $A=X$ y $B=V$, se obtiene:

- $\mathcal{M}(V : X) = \mathcal{M}(X) \oplus \mathcal{M}[(I - XX^-)V].$
- $r(V : X) = r(X) + r((I - XX^-)V) = r(X) + r(Z'V) = r(X) + r(VZ).$

entonces, se sigue que $r(VZ : X) = r(V : X).$

Sea $u = VZh = Vt$, entonces $\mathcal{M}(VZ) \subset \mathcal{M}(V)$, y por lo tanto se satisface, $\mathcal{M}(VZ : X) \subset \mathcal{M}(V : X)$ y como $r(VZ : X) = r(V : X)$, solo es posible que $\mathcal{M}(V : X) = \mathcal{M}(VZ : X).$ ■

Definición 18 Sean A y B matrices disjuntas con el mismo número de filas n . Cualquier vector $y \in \mathcal{M}(A : B)$ tiene solución única:

$$y = y_1 + y_2, \text{ donde } y_1 \in \mathcal{M}(A), y_2 \in \mathcal{M}(B)$$

entonces, $P_{A|B}$ se dice es un proyector sobre $\mathcal{M}(A)$ paralelo a (o a lo largo de) $\mathcal{M}(B)$ si y sólo si

$$P_{A|B}y = y_1 \text{ para todo } y \in \mathcal{M}(A : B).$$

Las propiedades de $P_{A|B}$ son puestas de manifiesto en los siguientes lemas.

Lema 19 Una condición necesaria y suficiente para que un operador $P_{A|B}$ satisfaga la definición anterior es

$$P_{A|B}A = A, P_{A|B}B = 0.$$

Demostración. De la Definición 18, si $y \in \mathcal{M}(A : B)$ podemos escribir

$$y = y_1 + y_2, \text{ donde } y_1 \in \mathcal{M}(A), y_2 \in \mathcal{M}(B)$$

o equivalentemente $y = Aa + Bb$, donde $y_1 = Aa \in \mathcal{M}(A)$, $y_2 = Bb \in \mathcal{M}(B)$. Luego si $P_{A|B}$ satisface la Definición 18, entonces $P_{A|B}$ es un proyector sobre $\mathcal{M}(A)$ paralelo a (o a lo largo de) $\mathcal{M}(B)$, si y sólo si $P_{A|B}y_1 = y_1$ para todo vector $y \in \mathcal{M}(A : B)$. Esto significa que:

$$\checkmark \quad P_{A|B}y_1 = P_{A|B}Aa = Aa = y_1$$

$$\checkmark \quad P_{A|B}y_2 = P_{A|B}Bb = 0$$

y dado que esto se cumple para todo $y \in \mathcal{M}(A : B)$, solo es posible que $P_{A|B}A = A$ y $P_{A|B}B = 0$.

Por otra parte, si $P_{A|B}A = A$ y $P_{A|B}B = 0$, se tiene que

$$P_{A|B}y = P_{A|B}(Aa + Bb) = P_{A|B}Aa + P_{A|B}Bb = Aa = y_1$$

Lema 20 $P_{A|B}$ es idempotente y única si $\mathcal{M}(A) + \mathcal{M}(B) = \mathcal{R}^n$. $\mathcal{R}^n$ el espacio euclidiano n -dimensional.

Lema 21 Una representación de $P_{A|B}$ es

$$P_{A|B} = A(C'A)^{-}C', \quad C = B^\perp.$$

En el caso cuando $B = A^\perp$, esto es, si A y B son complementos ortogonales que generan el espacio completo $\mathcal{M}(A) + \mathcal{M}(B) = \mathcal{R}^n$ se obtiene la matriz

$$P_{A|B} = A(C'A)^{-}C' = A(A'A)^{-}A'$$

que es el operador de proyección perpendicular sobre $\mathcal{M}(A)$, que se sabe es simétrica, idempotente y única para cualquier escogencia de la g-inversa de $A'A$. Es decir, si C es cualquier otra matriz que satisface la ecuación $\mathcal{M}(A) = \mathcal{M}(C)$, entonces $P_A = A(A'A)^{-}A' = C(C'C)^{-}C' = P_C$.

En el caso que B sea una matriz de máximo rango que satisfaga $A'\Sigma B = 0$, donde Σ es una matriz definida positiva (d.p.), del Lema 21 con $B^\perp = C = \Sigma A$ se sigue que $C' = A'\Sigma$ y entonces se obtiene al proyector

$$P_{A|B} = A(A'\Sigma A)^- A'\Sigma$$

al que en particular se denota con $P_{A\Sigma}$; esto es

$$P_{A\Sigma} = A(A'\Sigma A)^- A'\Sigma$$

De la expresión anterior se puede comprobar que $P_{A\Sigma}$ es idempotente, $\Sigma P_{A\Sigma}$ es simétrica y $P_{A\Sigma}$ es única para cualquier g-inversa $A'\Sigma A$. Además, si el producto interno se define como $(x, y) = x'\Sigma y$, el operador $P_{A\Sigma}$ es el proyector ortogonal sobre el espacio vectorial $\mathcal{M}(A)$.

A continuación se da la definición de proyector dentro de un espacio vectorial, que según Rao (1974, p.446) “fue introducida por Rao y Mitra (1971) y posteriormente examinada por Mitra y Rao”.

Definición 22 Sea M una matriz n.n.d. Una matriz P_{AM} se denomina un proyector dentro de $\mathcal{M}(A)$ con respecto a la norma o semi-norma definida por $\|x\| = (x'Mx)^{1/2}$, $x \in \mathcal{R}^n$ si y sólo si las siguientes dos condiciones se mantienen

$$P_{AM}Y \in \mathcal{M}(A), Y \in \mathcal{R}^n$$

$$\|Y - P_{AM}Y\| \leq \|Y - Ax\|, \forall x \in \mathcal{R}^m, Y \in \mathcal{R}^n.$$

De acuerdo a esta definición, la primera condición establece claramente que la proyección $P_{AM}y$ pertenece a $\mathcal{M}(A)$. La segunda condición señala que no existe un vector $Ax \in \mathcal{M}(A)$ diferente a $P_{AM}y$ cuya distancia a y sea menor a $\|y - P_{AM}y\|$ (medida con respecto a la norma o semi-norma $\|x\| = (x'Mx)^{1/2}$).

De acuerdo a Rao (1974), P_{AM} satisface las condiciones dadas en la definición anterior, si y sólo si satisface las siguientes dos condiciones:

$$\mathcal{M}(P_{AM}) \subset \mathcal{M}(A), \quad (2.13)$$

$$P'_{AM}MA = MA \quad (2.14)$$

de las cuales se desprende, que también cumple la condición

$$P_{AM}M = MP_{AM} \quad (2.15)$$

2.7. Las funciones lineales cero (LZF).

2.7.1. Definición

Según Bhimasankaram y Sengupta (1996), a partir de las propiedades de las LZF y argumentos sencillos de espacio vectorial, se pueden obtener los resultados principales en inferencia. Dado que las LZF son estimadores lineales insesgados de cero, basta con agregar una LZF a un estimador insesgado de $p'\beta$ para obtener otros estimadores lineales insesgados de $p'\beta$. Existen dos maneras en las que las LZF contribuyen a la inferencia. En la estimación de σ^2 , ya que contienen información relacionada con el error o ruido en el modelo, y en la estimación de la parte sistemática $X\beta$, pues la media y varianza de una LZF no depende de β , por lo que en cierto sentido están desligadas de $X\beta$, y por consiguiente pueden utilizarse para aislar el ruido que es útil para su estimación.

Definición 23 Una función lineal $l'y$ se denomina una función lineal cero (LZF) si $E(l'y) = 0$ para todos los posibles valores de β .

2.7.2. Relación de las funciones lineales cero con las funciones lineales estimables

Sea el modelo $y = X\beta + \varepsilon$ con $Cov(\varepsilon) = \sigma^2 V$, donde X y V pueden ser deficientes en rango. Un estimador insesgado de $X\beta$ es el mismo vector y , por lo tanto otro estimador insesgado para $X\beta$ se obtiene al sumar un vector de LZF $L'y$ al vector de respuestas y ; esto es, otro estimador insesgado para $X\beta$ es $y + L'y$ donde $L'y$ es un vector de funciones lineales cero cualesquiera.

De la definición de LZF se conoce que si $l'y$ es una LZF, $E(l'y) = l'X\beta = 0$ o equivalentemente que $X'l = 0$. Esto significa que los subespacios $\mathcal{M}(l)$ y $\mathcal{M}(X)$ son ortogonales y por lo tanto $\mathcal{M}(l) \subset \mathcal{M}(X)^\perp$, o equivalentemente que $\mathcal{M}(l) \subset \mathcal{M}(I - P_X)$.

Esta última relación implica que $l \in \mathcal{M}(I - P_X)$ y por tanto es posible expresar el vector l como $l = (I - P_X)m$ para algún vector m . En otras palabras, lo anterior

significa que las LZF son de la forma $l'y = m'(I - P_X)y = m'(I - P_X)(X\beta + \varepsilon) = m'(I - P_X)X\beta + m'(I - P_X)\varepsilon = m'(I - P_X)\varepsilon$; esto es, son función del vector de error ε del modelo y , por ello no dependen de β . De esta manera, cualquier estimador lineal insesgado de $X\beta$ es de la forma $y + L'y$, donde $L' = B(I - P_X)$ para alguna matriz B de $n \times n$.

Según Bhimasankaram y Sengupta (1996) la relación que enlaza las LZF con los BLUE de las funciones lineales paramétricas estimables $p'\beta$, está dada en el siguiente teorema, que refieren se debe a Rao C.R., según el cual el BLUE de $X\beta$ se corresponde con la escogencia de B para la cual el estimador $y + B(I - P_X)y$ es no correlacionado con $(I - P_X)y$.

Teorema 24 *En el modelo lineal $(y, X\beta, \sigma^2V)$ con V posiblemente singular, una función lineal $l'y$ es el BLUE de su valor esperado si y sólo si es no correlacionada con cada LZF.*

Come se muestra más adelante, las LZF permiten obtener una expresión particular del estimador de $X\beta$ para el modelo de Gauss-Markov $(y, X\beta, \sigma^2V)$ fundamental en esta investigación.

Capítulo 3

Marco Metodológico

En toda investigación, establecer el marco metodológico bajo el cual se llevará a cabo, es fundamental, puesto que define de manera clara la forma en la que el estudio se va a realizar. Así, en el presente trabajo, se toman en cuenta tres tipos de investigación: documental, exploratoria y descriptiva; considerándose como etapas que deben realizarse.

3.1. Tipos de Investigación

3.1.1. Investigación Documental

Según Risquez (1999): la investigación documental es aquella que tiene por propósito la revisión de fuentes documentales, recolectando, evaluando, verificando y sintetizando evidencias de lo que se investiga con el fin de establecer conclusiones relacionadas con los objetos de la investigación (p.40).

Esto le confiere a esta investigación el carácter de rigurosa, metódica y exhaustiva. La consulta y revisión de diversas fuentes de información secundaria como: documentos, artículos, informes de investigación, libros, tesis anteriores con temas relacionados, permitirá obtener información relevante y necesaria relacionada con el área de los modelos lineales. De esta manera, con la investigación documental se logrará alcanzar el nivel de conocimiento previo necesario, para avanzar satisfactoriamente hacia la búsqueda de

los objetivos del trabajo

3.1.2. Investigación Exploratoria

Según Hernández et. al. (1998) “Los estudios exploratorios tienen por objeto esencial familiarizarnos con un tópico desconocido o poco estudiado o novedoso” (p. 71).

En este caso, este trabajo posee características de la investigación exploratoria, al estudiar el modelo lineal en p-etapas (estudiado por pocos investigadores), desde el punto de vista de los operadores de proyección y las funciones lineales cero, con el objetivo de encontrar expresiones alternas a las existentes, para los estimadores del vector de parámetros y del vector de medias del modelo.

3.1.3. Investigación Descriptiva

Según Dankhe (1986) citado por Hernández et. al. (1998): “Los estudios descriptivos buscan especificar las propiedades importantes de personas, grupos, comunidades o cualquier otro fenómeno que sea sometido a análisis”. (p. 60). Durante el desarrollo de la tesis doctoral, se pretende aplicar los resultados obtenidos para el modelo en dos etapas, a una situación real. De esta manera, en su momento se hará una descripción de la población a estudiar y de las variables involucradas en el estudio. Se tomarán muestras aleatorias de datos, para cada variable involucrada y se realizará un análisis descriptivo de la situación, como parte de la investigación.

3.2. Diseño de la Investigación

Según Hernández et. al. (2006): “El término diseño se refiere al plan o estrategia concebida para obtener la información que se desea.... En el enfoque cuantitativo, el investigador utiliza su o sus diseños para aportar evidencia respecto de los lineamientos de la investigación”. (p.158)

En este sentido, el diseño de la investigación puede ser de dos tipos: Experimental y No Experimental. La propuesta que aquí se plantea se adecúa a los propósitos de

un diseño de investigación no experimental, porque no se manipulan variables, sólo se describe y diagnostica una problemática para determinar sus causas y ofrecer soluciones a la misma.

Hernández et. al. (1998) expresan que la investigación no experimental:

Es la que se realiza sin manipular deliberadamente variables. Es decir, se trata de la investigación donde no hacemos variar intencionalmente las variables independientes. Lo que hacemos en la investigación no experimental es observar fenómenos tal y como se dan en su contexto natural, para después analizarlos.
(p.184)

De acuerdo con lo anterior, las investigaciones no experimentales se identifican porque no se construyen situaciones, sólo se observan. Asimismo, los diseños no experimentales se pueden clasificar en transeccionales y longitudinales. En la aplicación del modelo en dos etapas, se pretende realizar un diseño transeccional, datos tomados en un solo momento, en su contexto natural. La investigación tendrá un diseño transeccional descriptivo, porque: “el procedimiento consiste en ubicar una o diversas variables a un grupo de personas u otros seres vivos, objetos, situaciones, contextos, fenómenos, comunidades; y así proporcionar su descripción”. (Hernández, 1998)

Capítulo 4

El Modelo Lineal en Dos Etapas Regular.

En este capítulo se dan algunos avances obtenidos en el desarrollo de la Tesis Doctoral. Específicamente se obtienen estimadores UBLUE para el vector de medias del modelo en dos etapas regular (1.1), y del modelo transformado (1.2), así como para el vector de parámetros β de ambos modelos.

Las expresiones obtenidas para β , son diferentes, pero equivalentes a las dadas por Volaufova (1987) para el estimador del vector de parámetros tanto del modelo transformado como del modelo lineal en dos etapas regular. En este trabajo se aplica el punto de vista de los operadores de proyección perpendicular de un espacio vectorial y se emplean los resultados dados por Rao (1974) y Bhimasankaran y Sengupta (1996), de manera que los estimadores obtenidos, son función de la proyección del vector de respuestas y sobre el espacio vectorial generado por la matriz X del modelo.

4.1. Resultados Preliminares

En esta sección se establece la relación entre algunos resultados dados por Rao (1974) y Bhimasankaran y Sengupta (1996) para el modelo $(y, X\beta, \sigma^2V)$, donde X puede ser deficiente en rango y V n.n.d. Además, se muestra la equivalencia de estos resultados con la expresión $\tilde{\beta} = (X'V_{\theta}^{-1}X)^{-1}X'V_{\theta}^{-1}y$ del estimador del parámetro β del modelo

en dos etapas empleada por Volaufova (1987).

4.1.1. Algunos Resultados de Interés.

Los siguientes lemas junto a la Definición 22, fueron empleados por Rao para derivar el BLUE de una función lineal paramétrica, en particular para obtener el BLUE del vector de medias del modelo $(y, X\beta, \sigma^2V)$, donde X puede ser deficiente en rango y V n.n.d. Las demostraciones son las dadas por Rao en su artículo, solo que en este trabajo, han sido desarrolladas en detalle por el autor, incluyéndose algunos pasos intermedios.

Lema 25 *Una representación de P_{AM} es*

$$P_{AM} = A(A'MA)^-A'M \quad (4.1)$$

Demostración. La prueba consiste en verificar que (4.1) satisface $\mathcal{M}(P_{AM}) \subset \mathcal{M}(A)$ y $P'_{AM}MA = MA$, y en consecuencia cumple con las condiciones dadas en la Definición 22. $\mathcal{M}(P_{AM}) \subset \mathcal{M}(A)$ se sigue del hecho que $P_{AM} = A(A'MA)^-A'M = AH$ para alguna matriz H .

Para probar la condición $P'_{AM}MA = MA$, descomponemos $M = QQ'$ aplicando el Teorema B.22 dado por Christensen (2011, p.424), ya que según la definición 22 M es n.n.d. y por tanto simétrica, obteniéndose al sustituir en la expresión para P_{AM} :

$$\begin{aligned} P'_{AM}MA &= MA(A'MA)^-A'MA \\ &= QQ'A(A'QQ'A)^-A'QQ'A \\ &= Q[Q'A]([Q'A]'[Q'A])^-[Q'A]'[Q'A] \\ &= QP_{Q'A}Q'A \\ &= QQ'A \\ &= MA \end{aligned} \quad (4.2)$$

al ser $P_{Q'A} = [Q'A]([Q'A]'[Q'A])^-[Q'A]'$ el operador de proyección perpendicular sobre $\mathcal{M}(Q'A)$, y por lo tanto $P_{Q'A}Q'A = Q'A$. ■

El siguiente lema $\{P_{AM}\}$ denota la clase de todas las matrices que satisfacen las condiciones (2.13) y (2.14).

Lema 26 Sean M y A como en la Definición 22 y denótese $A^\perp$ por H . Además, sea $T = M + AUA'$ donde U es cualquier matriz simétrica tal que $R(T) = R(M : A)$ y T n.n.d. Denótese por T^- cualquier g -inversa n.n.d. de T . Entonces:

$$(i) \quad \{P'_{AM}\} = \{P_{MA|H}\}$$

$$(ii) \quad P'_{HM} + P_{A|MH} \equiv I \text{ en } \mathcal{M}(M : A)$$

$$(iii) \quad P_{A|MH} \equiv P_{AT^-} \text{ en } \mathcal{M}(M : A)$$

Demostración. Prueba de (i): Del Lema 17 con $X = A^\perp = H$, $Z = A$, y $V = M$, se sigue que H y MA son disjuntos. Al aplicar el Lema 21, con la A del lema igual a MA , $B = H$ y $C = B^\perp = H^\perp = A$, se obtiene $P_{MA|H} = MA(C'MA)^-C' = MA(A'MA)^-A'$. Luego $P'_{MA|H} = A(A'MA)^-A'M$, la cual es la expresión para P_{AM} dada en el Lema 25, luego $P'_{MA|H} = P_{AM}$. En consecuencia, $\mathcal{M}(P'_{MA|H}) = \mathcal{M}(P_{AM}) \subset \mathcal{M}(A)$ y por lo tanto se satisface la condición (2.13). Por otra parte, del Teorema B.22 de Christensen (2011, p.424), se puede escribir $M = QQ'$. De esta manera se obtiene:

$$\begin{aligned} P_{MA|H}MA &= [MA(A'MA)^-A'] [MA] = MA(A'MA)^-A'MA \\ &= QQ'A(A'QQ'A)^-A'QQ'A \\ &= Q[Q'A]([Q'A]'[Q'A])^-[Q'A][Q'A] \\ &= Q[Q'A] = QQ'A \\ &= MA \end{aligned}$$

satisfaciéndose así la condición (2.14), y así $\{P'_{AM}\} = \{P_{MA|H}\}$.

Prueba de (ii): Dado que M es n.n.d. y $H = A^\perp$, del Lema 17 con $X = A$ y $V = M$ se tiene que A y MH son disjuntos y que $\mathcal{M}(M : A) = \mathcal{M}(MH : A)$. Entonces se puede escribir, $\mathcal{M}(M : A) = \mathcal{M}(MH : A) = \mathcal{M}(MH) \oplus \mathcal{M}(A)$.

Sea y un vector cualquiera tal que $y \in \mathcal{M}(M : A) = \mathcal{M}(MH : A) = \mathcal{M}(A) \oplus \mathcal{M}(MH)$, se sigue que podemos escribir y de manera única como

$$y = Aa + MHb$$

con $Aa \in \mathcal{M}(A)$ y $MHb \in \mathcal{M}(MH)$. Como A y MH son disjuntas, de la Definición 18, se sigue que para $y \in \mathcal{M}(M : A) = \mathcal{M}(MH : A)$; $P_{A|MH}$ es un proyector sobre $\mathcal{M}(A)$ a lo largo de $\mathcal{M}(MH)$ y $P_{MH|A}$ un proyector sobre $\mathcal{M}(MH)$ a lo largo de $\mathcal{M}(A)$.

Por las propiedades de los proyectores se tiene que:

$$\begin{aligned} [P_{MH|A} + P_{A|MH}]y &= P_{MH|A}y + P_{A|MH}y \\ &= P_{MH|A}[Aa + MHb] + P_{A|MH}[Aa + MHb] \\ &= MHb + Aa = y \\ &= Iy \end{aligned}$$

Como $[P_{MH|A} + P_{A|MH}]y = Iy$ para cualquier vector $y \in \mathcal{M}(M : A)$, solo es posible que

$$P_{MH|A} + P_{A|MH} = I \text{ en } \mathcal{M}(M : A)$$

Aplicando (i), se tiene $P'_{HM} = P_{MH|A}$, y al sustituir en la ecuación anterior se obtiene:

$$P'_{HM} + P_{A|MH} = I \text{ en } \mathcal{M}(M : A).$$

Prueba de (iii): En este caso, debe probarse que en $\mathcal{M}(M : A)$, los operadores $P_{A|MH}$ y P_{AT-} tienen las mismas propiedades. Del Lema 19 se cumple:

$$P_{A|MH}A = A, P_{A|MH}MH = 0$$

entonces, lo que debe probarse es que P_{AT-} también satisface:

$$P_{AT-}A = A, P_{AT-}MH = 0.$$

Sea $P_{AT^-} = A(A'T^-A)^-A'T^-$. Como T^- es n.n.d., a través de un procedimiento similar al utilizado en la demostración del Lema 25, se puede mostrar que P_{AT^-} satisface las condiciones (2.13) a (2.15). De aquí, se cumple:

$$(a) \mathcal{M}(P_{AT^-}) \subset \mathcal{M}(A)$$

$$(b) P'_{AT^-}T^-A = T^-A$$

$$(c) P'_{AT^-}T^- = T^-P_{AT^-}$$

De (a) se sigue que $\mathcal{M}(P_{AT^-}) \subset \mathcal{M}(A) \subset \mathcal{M}(M : A) = \mathcal{M}(T) = \mathcal{M}(TT^-)$, luego $TT^-A = A$ y $TT^-P_{AT^-} = P_{AT^-}$. A partir de la ecuación (b) se obtiene $TP'_{AT^-}T^-A = TT^-A = A$, y de la ecuación (c), luego de pre multiplicarla por T y pos multiplicarla por A , se consigue $TP'_{AT^-}T^-A = P_{AT^-}A$. Entonces, al comparar las dos últimas ecuaciones, se aprecia que $TP'_{AT^-}T^-A = P_{AT^-}A = A$.

Por ser T^- n.n.d. se cumple que $(T^-)' = T^-$. Entonces de la ecuación (c) anterior se obtiene:

$$\begin{aligned} P'_{AT^-}T^- &= T^-P_{AT^-} \\ TP'_{AT^-}T^- &= TT^-P_{AT^-} = P_{AT^-} \\ TP'_{AT^-}T^-T &= P_{AT^-}T \\ TP'_{AT^-} &= P_{AT^-}T \\ TP'_{AT^-}H &= P_{AT^-}TH = P_{AT^-}MH = 0 \end{aligned}$$

ya que al ser $A'H = 0$, se sigue que $TH = (M + AUA')H = MH$ y $P'_{AT^-}H = T^-A(A'T^-A)^-A'H = 0$. Dado que $P_{AT^-}A = A$ y $P_{AT^-}MH = 0$, se concluye que $P_{A|MH} \equiv P_{AT^-}$. ■

4.1.2. Caracterización del BLUE de $X\beta$

Sea el modelo lineal general $(y, X\beta, \sigma^2V)$, donde V puede ser singular y X deficiente en rango. Según Rao (1974), el BLUE de una función lineal paramétrica $p'\beta$ es la función

lineal L'_*y si se cumplen las dos propiedades siguientes:

$$L'_*X = p' \quad (4.3)$$

$$L'_*VL_* \leq L'VL \forall L \text{ tal que } L'X = p' \quad (4.4)$$

Rao (1974) en los Teoremas 3.1 y 3.2 de su artículo caracteriza las formas funcionales del BLUE de una función paramétrica $p'\beta$ y del vector de medias $X\beta$ respectivamente. Específicamente, Rao da los siguientes enunciados.

Teorema 27 *Sea L_0 cualquier vector tal que $L'_0X = p'$. Entonces L'_*y es el BLUE de $p'\beta$, donde*

$$L_* = (I - P_{ZV})L_0 = (I - P'_{VZ|X})L_0$$

Demostración. Lo que se debe probar es que L_* satisface $L'_*VL_* \leq L'VL \forall L$ tal que $L'X = p'$. Sea $Z = X^\perp$ y $L = L_0 - Z\lambda$ una solución para $L'X = p'$, ya que $(L'_0 - \lambda'Z')X = L'_0X = p'$, donde λ es un vector cualquiera de dimensión apropiada.

Dado que V es n.n.d., al aplicar la Definición 22, con $M = V$, $A = Z$, $Y = L_0$ y $X = \lambda$ se obtiene:

$$\|L_0 - P_{ZV}L_0\|_V \leq \|L_0 - Z\lambda\|_V,$$

donde $P_{ZV} = Z(Z'VZ)^-Z'V$ según el Lema 25. Luego se cumple que:

$(L_0 - P_{ZV}L_0)'V(L_0 - P_{ZV}L_0) \leq (L_0 - Z\lambda)'V(L_0 - Z\lambda)$. Denominando $L_* = L_0 - P_{ZV}L_0$ y sustituyendolo en la desigualdad anterior, se cumple:

$$L'_*VL_* \leq L'VL \forall L \text{ tal que } L'X = p'$$

en consecuencia, L_* cumple con las condiciones (4.3) y (4.4) y por lo tanto $L'_*y = [(I - P_{ZV})L_0]'y$ es el BLUE de $p'\beta$.

Solo falta probar que $(I - P_{ZV})L_0 = (I - P'_{VZ|X})L_0$. Para ello observe de la parte (i) del Lema 26 que $\{P'_{ZV}\} = \{P_{VZ|X}\}$, o equivalentemente que $P_{ZV} = P'_{VZ|X}$, y al sustituir en la expresión para el BLUE de $p'\beta$ se obtiene $L'_*y = [(I - P_{ZV})L_0]'y = [(I - P'_{VZ|X})L_0]'y$. ■

Corolario 28 Sea Λ_* una matriz $n \times k$ tal que $\Lambda'_0 X = \pi'$. Entonces $\Lambda'_* y$ es el BLUE de $\pi' \beta$ (k funciones paramétricas) donde

$$\Lambda_* = (I - P_{ZV}) \Lambda_0 = \left(I - P'_{VZ|X} \right) \Lambda_0$$

Demostración. Basta con sustituir Λ_0 por L_0 en la expresión de L_* dada en el Teorema 27. ■

Teorema 29 El BLUE de $X\beta$ puede ser expresado en las formas alternativas

$$(a) (I - P'_{ZV})y,$$

$$(b) P_{X|VZ}y = (I - P_{VZ|X})y,$$

$$(c) P_{XT^-}y.$$

En (c), T^- es una g -inversa n.n.d. de $T = V + XU'X'$ para cualquier matriz simétrica U tal que T es n.n.d. y $r(T) = r(V : X)$.

Demostración. Prueba de (a). En este caso, basta con hacer $\Lambda_0 = I$ en el Corolario 28 para que $\pi' \beta = X\beta$. Entonces el BLUE de $X\beta$ estará dado por $\Lambda'_* y = (I - P'_{ZV})y = (I - P_{VZ|X})y$.

Prueba de (b). Del Lema 2.3 de Rao (1974) se sabe que $y \in \mathcal{M}(V : X)$ y de la parte (ii) del Lema 26 con $X = A$, $M = V$ y $H = A^\perp = X^\perp = Z$, se tiene que $P'_{ZV} + P_{X|VZ} \equiv I$ en $\mathcal{M}(V : X)$. Al despejar P'_{ZV} de esta ecuación y sustituirla en la expresión del BLUE $\Lambda'_* y$ dada en (a) se obtiene que $\Lambda'_* y = (I - P'_{ZV})y = (I - P_{VZ|X})y = P_{X|VZ}y$.

Prueba de (c). Este resultado es una consecuencia directa de la aplicación de la parte (iii) del Lema 26 con $X = A$, $M = V$ y $H = A^\perp = X^\perp = Z$, obteniéndose como resultando que $P_{X|VZ} \equiv P_{XT^-}$ en $\mathcal{M}(V : X)$. Luego el BLUE de $X\beta$ también puede escribirse como $\Lambda'_* y = P_{X|VZ}y = P_{XT^-}y$. ■

Del Lema 25 y el teorema anterior se sigue que el BLUE de $X\beta$, para el modelo $(y, X\beta, \sigma^2 V)$, donde X puede ser deficiente en rango y V n.n.d. puede expresarse en

cualquiera de las siguientes formas equivalentes:

$$\widehat{X\beta} = [I - P'_{ZV}]y = [I - VZ(Z'VZ)^{-1}Z']y \quad (4.5)$$

$$= P_{XT^-}y = X(X'T^-X)^{-1}X'T^-y. \quad (4.6)$$

4.1.3. Las funciones lineales cero y el BLUE de $X\beta$

Veintidós años después de haber sido publicado el artículo de Rao sobre proyectores, Bhimasankaran y Sengupta (1996) publicaron un trabajo donde relacionan a las funciones lineales cero (LZF) con las funciones lineales paramétricas estimables, para el modelo $(y, X\beta, \sigma^2V)$. Ellos utilizaron el Teorema 24 junto con el siguiente lema para deducir la expresión del BLUE del vector de medias $X\beta$ del modelo. En este lema a la matriz de covarianza del vector v se le denota con V_v .

Lema 30 Sea $z = (u' : v')'$ un vector aleatorio que tiene momentos de primer y segundo orden tal que $E(v) \in \mathcal{M}(V_v)$. Entonces la combinación lineal $u + Bv$ es no correlacionada con v si y solo si $Bv = -Cov(u, v)[V_v]^{-1}v$.

Demostración. Si $E(v) \in \mathcal{M}(V_v)$, se puede escribir $E(v) = V_v h$ para algún h . Entonces se tiene

$$E[(I - P_{V_v})v] = (I - P_{V_v})E[v] = (I - P_{V_v})V_v h = 0$$

$$Cov((I - P_{V_v})v) = (I - P_{V_v})Cov(v)(I - P_{V_v}) = (I - P_{V_v})V_v(I - P_{V_v}) = 0$$

luego, $Pr[(I - P_{V_v})v = 0] = 1$ o equivalentemente $Pr[P_{V_v}v = v] = 1$, en consecuencia se obtiene $v \in \mathcal{M}(P_{V_v}) = \mathcal{M}(V_v)$

Como $Cov(u + Bv, v) = Cov(u, v) + BCov(v) = Cov(u, v) + BV_v$. Para que la combinación lineal $u + Bv$ y el vector v sean no correlacionados, debe ocurrir que

$Cov(u, v) + BV_v = 0$, y de esta manera,

$$\begin{aligned} BV_v &= -Cov(u, v) \\ B[V_v][V_v]^{-1}v &= -Cov(u, v)[V_v]^{-1}v \\ Bv &= -Cov(u, v)[V_v]^{-1}v \end{aligned}$$

■

De la sección 2.7.2 se sabe que las LZF son de la forma $l'y = m'(I - P_X)y$. Como $E(y) = X\beta$, Bhimasankaran y Sengupta (1996) señalan que cualquier estimador lineal insesgado de $X\beta$ debe ser de la forma $y + B(I - P_X)y$ para alguna matriz B . De esta manera, de acuerdo al Teorema 24, el BLUE del vector de medias $X\beta$ se obtiene al seleccionar la matriz B de tal manera que $y + B(I - P_X)y$ sea no correlacionado con $(I - P_X)y$.

En este sentido, del Lema 30 se conoce que $Cov(u + Bv, v) = 0$ si y solo si $Bv = -Cov(u, v)[Cov(v)]^{-1}v$. Por lo tanto, escogiendo $u = y$ y $v = (I - P_X)y$, se sigue que $y + B(I - P_X)y$ es no correlacionado con $(I - P_X)y$ si y sólo si $Bv = B(I - P_X)y = -Cov(y, (I - P_X)y)\{Cov[(I - P_X)y]\}^{-1}(I - P_X)y$.

Como $Cov(y, (I - P_X)y) = Cov(y)(I - P_X) = \sigma^2 V(I - P_X)$ y $Cov[(I - P_X)y] = (I - P_X)\sigma^2 V(I - P_X)$, del Teorema 24 se sigue, que el BLUE de $X\beta$ esta dado por:

$$\begin{aligned} \widehat{X\beta} &= y + B(I - P_X)y \\ \widehat{X\beta} &= y - \sigma^2 V(I - P_X)\{(I - P_X)\sigma^2 V(I - P_X)\}^{-1}(I - P_X)y \\ \widehat{X\beta} &= [I - V(I - P_X)\{(I - P_X)V(I - P_X)\}^{-1}(I - P_X)]y \end{aligned} \quad (4.7)$$

Dado que $X'(I - P_X) = X' - X' = 0$ y $r(I - P_X) = n - r(P_X) = n - r(X)$, la matriz $I - P_X$ es un caso particular de una matriz de máximo rango $Z = X^\perp$ tal que $X'Z = 0$. Si denominamos $Z = I - P_X$, la ecuación (4.7) podemos escribirla como

$$\widehat{X\beta} = [I - VZ(Z'VZ)^{-1}Z']y = [I - P'_{ZV}]y \quad (4.8)$$

Obsérvese que la ecuación (4.8) es igual a la ecuación (4.5) dada por Rao (1974). En consecuencia, el estimador BLUE de $X\beta$ dado por Bhimasankaran y Sengupta (1996) es un caso particular del obtenido por Rao (1974), cuando $Z = I - P_X$.

4.1.4. El estimador de $X\beta$ cuando V es no negativa definida

Si en la expresión (4.7) hacemos $V = I$, esta simplifica en

$$\begin{aligned}\widehat{X\beta} &= [I - V(I - P_X)\{(I - P_X)V(I - P_X)\}^-(I - P_X)]y \\ &= [I - (I - P_X)\{(I - P_X)(I - P_X)\}^-(I - P_X)]y \\ &= [I - (I - P_X)(I - P_X)^-(I - P_X)]y \\ &= [I - (I - P_X)]y = P_X y \\ &= X(X'X)^-X'y\end{aligned}$$

la cual es la expresión del BLUE de $X\beta$ para el modelo homocedástico $(y, X\beta, \sigma^2 V)$, donde $X(X'X)^-X'$ es la matriz operador de proyección perpendicular sobre $\mathcal{M}(X)$.

En su artículo Bhimasankaram y Sengupta (1996) también señalan que si U es una matriz simétrica para la cual $W = V + XUX'$ satisface $\mathcal{M}(W) = \mathcal{M}(V : X)$, la matriz V en (4.7) puede ser reemplazada con W al ser $X'(I - P_X) = 0$, para luego a partir de la factorización de rango de W como CC' obtener que $CC^{-L}(X : y) = (X : y)$ con probabilidad 1 y por lo tanto el BLUE de $X\beta$ estaría dado por

$$\widehat{X\beta} = X(X'W^-X)^-X'W^-y \quad (4.9)$$

con probabilidad 1.

Lo interesante del resultado dado en la expresión (4.9) es que es exactamente igual al mostrado en (4.6) con $W = T$, el cual fue dado por Rao (1974) en la parte (c) del Teorema 3.2 de su artículo.

Para una mayor comprensión de la relación entre las ecuaciones (4.7) y (4.9), a continuación se da la demostración realizada por el autor de este trabajo, donde se muestra que estas ecuaciones son equivalentes, por lo que es indiferente cual de ellas se

utilice para obtener el BLUE de $X\beta$ del modelo lineal general $(y, X\beta, \sigma^2V)$, donde V puede ser singular y X deficiente en rango.

Demostración. Supóngase que U es una matriz simétrica tal que $W = V + XUX'$ satisface $\mathcal{M}(W) = \mathcal{M}(V : X)$. Entonces se cumple $\mathcal{M}(V) \subset \mathcal{M}(W)$ y $\mathcal{M}(X) \subset \mathcal{M}(W)$. Al sustituir $V = W - XUX'$ en (4.7), sabiendo que $X'(I - P_X) = 0$ se obtiene

$$\widehat{X\beta} = [I - W(I - P_X)\{(I - P_X)W(I - P_X)\}^-(I - P_X)]y \quad (4.10)$$

Por ser U, V simétricas, W es simétrica n.n.d. y del Teorema 14.3.7 se puede descomponer en $W = CC'$, con C de rango completo por columnas, de tal modo que si $r(W) = k$, entonces $r(C) = r(C') = k$. Además de la Proposición B.51 de Christensen (2011) se sigue que $\mathcal{M}(C) = \mathcal{M}(CC') = \mathcal{M}(W) = \mathcal{M}(V : X)$. De esta manera se cumple:

1. $P_C = C(C'C)^{-1}C'$ es la m.p.p sobre $\mathcal{M}(W) = \mathcal{M}(X : V)$.
2. $P_C X = X$ y $P_C V = V$
3. Escogiendo $C^{-L} = (C'C)^{-1}C'$, $P_C = CC^{-L}$ y entonces $CC^{-L}X = X$, $CC^{-L}y = y$ ya que $\mathcal{M}(X) \subset \mathcal{M}(W) = \mathcal{M}(C)$ e $y \in \mathcal{M}(X : V) = \mathcal{M}(W)$, lo que implica $CC^{-L}(X : y) = (X : y)$.

donde a $C^{-L} = (C'C)^{-1}C'$ se le conoce como la matriz inversa por la izquierda de C .

Con lo anterior en mente, al sustituir $W = CC'$ e $y = CC^{-L}y$ en (4.10) se obtiene

$$\begin{aligned} \widehat{X\beta} &= [I - CC'(I - P_X)\{(I - P_X)CC'(I - P_X)\}^-(I - P_X)]CC^{-L}y \\ &= [C - CC'(I - P_X)\{(I - P_X)CC'(I - P_X)\}^-(I - P_X)C]C^{-L}y \\ &= C[I_k - C'(I - P_X)\{(I - P_X)CC'(I - P_X)\}^-(I - P_X)C]C^{-L}y \\ &= C[I_k - [C'(I - P_X)]\{[C'(I - P_X)]'[C'(I - P_X)]\}^{-1}[C'(I - P_X)]'C]C^{-L}y \\ &= C[I_k - P_{C'(I - P_X)}]C^{-L}y \end{aligned} \quad (4.11)$$

donde $P_{C'(I - P_X)}$ e $I_k - P_{C'(I - P_X)}$ son las matrices de proyección perpendicular sobre los espacios vectoriales $\mathcal{M}(C'(I - P_X))$ y $\mathcal{M}(C'(I - P_X))^\perp$ respectivamente.

Obsérvese que

$$(C^{-L}X)'C'(I-P_X) = X'(C^{-L})'C'(I-P_X) = X'C(C'C)^{-1}C'(I-P_X) = X'(I-P_X) = 0$$

lo que significa $\mathcal{M}(C^{-L}X) \perp \mathcal{M}(C'(I-P_X))$, o equivalentemente que $\mathcal{M}(C^{-L}X) \subset \mathcal{M}(C'(I-P_X))^\perp$. En consecuencia, para probar que $\mathcal{M}(C^{-L}X) = \mathcal{M}(C'(I-P_X))^\perp$, es suficiente probar que

$$r[\mathcal{M}(C^{-L}X)] = r[\mathcal{M}(C'(I-P_X))^\perp].$$

Dado que una representación de la g-inversa de X es $X^- = (X'X)^-X'$ se sigue que $\mathcal{M}((I-XX^-)C) = \mathcal{M}((I-X(X'X)^-X')C) = \mathcal{M}((I-P_X)C)$. Por otra parte, como $\mathcal{M}(X) \subset \mathcal{M}(C)$, se verifica que $\mathcal{M}(C:X) = \mathcal{M}(C)$, y $r(C:X) = r(C)$, por lo que según el corolario 17.2.9 de Harville (1997) se obtiene:

$$\mathcal{M}(C:X) = \mathcal{M}(X) \oplus \mathcal{M}((I-XX^-)C) = \mathcal{M}(X) \oplus \mathcal{M}((I-P_X)C)$$

$$r(C:X) = r(X) + r((I-XX^-)C) = r(X) + r((I-P_X)C) = r(C)$$

luego

$$r((I-P_X)C) = r(C) - r(X) \Rightarrow r(C'(I-P_X)) = r(C) - r(X) = k - r(X).$$

Lo ecuación anterior significa que $r[\mathcal{M}(C'(I-P_X))] = k - r(X)$ o equivalentemente que $r[\mathcal{M}(C'(I-P_X))^\perp] = r(X)$. Como $CC^{-L}X = X$ y C de rango completo por columnas, se sigue al aplicar el Lema 8.3.2 del Harville que

$$r(CC^{-L}X) = r(C^{-L}X) = r(X) \Rightarrow r[\mathcal{M}(C^{-L}X)] = r(X)$$

y como que $\mathcal{M}(C^{-L}X) \subset \mathcal{M}(C'(I-P_X))^\perp$ solo es posible que

$$\mathcal{M}(C^{-L}X) = \mathcal{M}(C'(I-P_X))^\perp.$$

Dado que las matrices operador de proyección perpendicular son únicas, se sigue que $I_k - P_{C'(I-P_X)} = P_{C^{-L}X}$. Al sustituir en la ecuación (4.11) y recordando que $C^{-L} = (C'C)^{-1}C'$ se obtiene

$$\begin{aligned} \widehat{X\beta} &= C(I_k - P_{C'(I-P_X)})C^{-L}y \\ &= CP_{C^{-L}X}C^{-L}y \\ &= C\{C^{-L}X[(C^{-L}X)'(C^{-L}X)]^-(C^{-L}X)'\}C^{-L}y \\ &= C(C'C)^{-1}C'X[X'C(C'C)^{-1}(C'C)^{-1}C'X]^-X'C(C'C)^{-1}(C'C)^{-1}C'y \end{aligned}$$

Como $\mathcal{M}(W) = \mathcal{M}(C) = \mathcal{M}(X : V)$ e $y \in \mathcal{M}(X : V)$, se cumple:

$$1) \quad WW^{-}X = X \Rightarrow X' = X'W^{-}W$$

$$2) \quad C(C'C)^{-1}C'X = P_C X = X$$

$$3) \quad C(C'C)^{-1}C'y = P_C y = y$$

en consecuencia,

$$\begin{aligned} \widehat{X\beta} &= C(C'C)^{-1}C'X[X'C(C'C)^{-1}(C'C)^{-1}C'X]^{-}X'C(C'C)^{-1}(C'C)^{-1}C'y \\ &= X[X'C(C'C)^{-1}(C'C)^{-1}C'X]^{-}X'C(C'C)^{-1}(C'C)^{-1}C'y \\ &= X[X'W^{-}WC(C'C)^{-1}(C'C)^{-1}C'WW^{-}X]^{-}X'W^{-}WC(C'C)^{-1}(C'C)^{-1}C'y \end{aligned}$$

y como $W = CC'$, se obtiene

$$\begin{aligned} \widehat{X\beta} &= X[X'W^{-}CC'C(C'C)^{-1}(C'C)^{-1}C'CC'W^{-}X]^{-}X'W^{-}CC'C(C'C)^{-1}(C'C)^{-1}C'y \\ &= X[X'W^{-}CC'W^{-}X]^{-}X'W^{-}C(C'C)^{-1}C'y \\ &= X[X'W^{-}WW^{-}X]^{-}X'W^{-}y \\ &= X[X'W^{-}X]^{-}X'W^{-}y \end{aligned}$$

obteniéndose la expresión dada en (4.9). Obsérvese que esta expresión es idéntica a la ecuación (4.6) con $T = W$, y por lo tanto se tiene que

$$\widehat{X\beta} = X[X'W^{-}X]^{-}X'W^{-}y = P_{XW^{-}y} \quad (4.12)$$

la cual es la expresión dada por Rao (1974) en la parte (c) del Teorema 3.2 de su artículo. ■

Según Bhimasankaran y Sengupta (1996) un método alternativo para obtener la expresión (4.12) es el de la matriz particionada de Rao (1973). En su artículo, ellos también señalan que cuando la matriz V es definida positiva, la matriz W en la expresión (4.12) puede sustituirse por V para obtener la expresión convencional del BLUE de $X\beta$

dada por

$$\widehat{X\beta} = X (X'V^{-1}X)^{-} X'V^{-1}y$$

que en el caso en que la matriz X sea de rango completo por columnas, el único estimador del parámetro β está dado por:

$$\widehat{\beta} = (X'V^{-1}X)^{-} X'V^{-1}y$$

la cual es igual a la ecuación dada en (2.12) e igual a la empleada por Volaufova (1987) para encontrar el estimador de β para el modelo en dos etapas regular.

4.2. Los modelos lineales en dos etapas

En la Sección 1.3 de este trabajo se hizo mención de los modelos en dos etapas y de un modelo al que se denominó el modelo transformado. Sin embargo, dada la importancia de los mismos en esta investigación, a continuación se consideran nuevamente sus definiciones.

Sea

$$y_1 = X_1\beta_1 + \varepsilon_1;$$

$$y_2 = D\beta_1 + X_2\beta_2 + \varepsilon_2$$

un modelo donde y_1 y y_2 son vectores aleatorios de dimensión $n_1 \times 1$, $n_2 \times 1$ respectivamente; las matrices X_1 , X_2 , y D conocidas de números reales; β_1 , β_2 vectores de parámetros desconocidos de dimensión $p_1 \times 1$, $p_2 \times 1$; y ε_1 , ε_2 vectores de error aleatorios con $E(\varepsilon_1) = 0$; $Cov(\varepsilon_1) = \sigma_1^2 V_1$, $E(\varepsilon_2) = 0$; $Cov(\varepsilon_2) = \sigma_2^2 V_2$, ε_1 y ε_2 no correlacionados, $\sigma_1^2 > 0$, $\sigma_2^2 > 0$ parámetros desconocidos y V_1 , V_2 matrices conocidas de dimensión apropiada. A este modelo se le denomina modelo lineal en dos etapas.

Cuando en el modelo en dos etapas los rangos de las matrices X_1 , X_2 , V_1 , V_2 son $r(X_1) = p_1 < n_1$, $r(X_2) = p_2 < n_2$, $r(V_1) = n_1$, $r(V_2) = n_2$, Kubáček (1988) lo denomina

modelo lineal en dos etapas regular, y puede escribirse matricialmente como

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ D & X_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix} \quad (4.13)$$

$$y = X\beta + \varepsilon$$

con

$$Cov(\varepsilon) = \begin{bmatrix} \sigma_1^2 V_1 & 0 \\ 0 & \sigma_2^2 V_2 \end{bmatrix} = V_\varepsilon,$$

En otras palabras, en el modelo lineal en dos etapas regular las matrices X_1, X_2 son de rango completo por columnas, V_1, V_2 son definidas positivas, y $\sigma_1^2 > 0, \sigma_2^2 > 0$ parámetros de covarianza desconocidos.

Si y en el modelo (4.13) es pre-multiplicado por la matriz F que sigue:

$$F = \begin{bmatrix} I_1 & 0 \\ -DQ & I_2 \end{bmatrix}$$

se obtiene el modelo transformado $Fy = FX\beta + F\varepsilon$, que en forma matricial es

$$\begin{bmatrix} y_1 \\ y_2^* \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2^* \end{bmatrix} \quad (4.14)$$

$$y^* = X^*\beta + \varepsilon^*$$

con

$$Cov(\varepsilon^*) = V_{\varepsilon^*} = \begin{bmatrix} \sigma_1^2 V_1 & -\sigma_1^2 C' \\ -\sigma_1^2 C & \sigma_1^2 C V_1^{-1} C' + \sigma_2^2 V_2 \end{bmatrix}$$

en donde $y^* = Fy$, $X^* = FX$, $\varepsilon^* = F\varepsilon$, $y_2^* = -DQy_1 + y_2$, $\varepsilon_2^* = -DQ\varepsilon_1 + \varepsilon_2$, $C = DQV_1$ y Q una matriz tal que $QX_1 = I$.

Los resultados a continuación, enteramente desarrollados por el autor, corresponden a la primera parte del trabajo de su tesis doctoral. En la sección 4.3 se obtiene el

UBLUE para el vector de medias $X\beta$ del modelo transformado, asumiendo que la razón $\rho = \sigma_2^2/\sigma_1^2$ es conocida. Posteriormente en la sección 4.4 se obtienen los UBLUE del vector de parámetros β y del vector de medias $X\beta$ del modelo en dos etapas regular, manteniéndose el supuesto sobre ρ . En la sección 4.5 los resultados obtenidos en las secciones 4.3 y 4.4 junto a otros resultados para el caso general, cuando $\sigma_1^2 > 0, \sigma_2^2 > 0$, se resumen en un teorema, el cual es formalmente demostrado. Se desea destacar que todos estos resultados son obtenidos a partir de las teorías sobre proyectores desarrolladas por Rao (1974) y Bhimasankaran y Sengupta (1996) que involucran proyectores oblicuos.

Las ventajas de la metodología empleada residen en 1) la obtención de expresiones equivalentes a las existentes, para los estimadores del vector de medias y del vector de parámetros del modelo en dos etapas en la forma tradicional; esto es, como el producto Py , donde P es un proyector oblicuo sobre el espacio $\mathcal{M}(X)$, generado por las columnas de la matriz X , denominado por Bhimasankaran y Sengupta (1996) como el espacio de estimación e y el vector de respuestas de la variable independiente de interés, 2) permite generalizar los resultados obtenidos al caso de un modelo regular con p etapas y 3) generalizar los resultados a la situación más general, donde tanto la matriz de diseño X del modelo como la matriz de covarianzas pueden ser deficientes en rango (a desarrollar por el autor como parte de su línea de investigación posterior a la tesis doctoral).

4.3. UBLUE para $X^*\beta$ cuando $\rho = \sigma_2^2/\sigma_1^2$ es conocido

Sea el modelo transformado (4.14), el cual se reescribe a continuación

$$\begin{bmatrix} y_1 \\ y_2^* \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2^* \end{bmatrix} \quad (4.15)$$

$$y^* = X^*\beta + \varepsilon^*$$

con

$$Cov(\varepsilon^*) = V_{\varepsilon^*} = \begin{bmatrix} \sigma_1^2 V_1 & -\sigma_1^2 C' \\ -\sigma_1^2 C & \sigma_1^2 C V_1^{-1} C' + \sigma_2^2 V_2 \end{bmatrix}$$

en donde $y^* = Fy$, $X^* = FX$, $\varepsilon^* = F\varepsilon$, $y_2^* = -DQy_1 + y_2$, $\varepsilon_2^* = -DQ\varepsilon_1 + \varepsilon_2$, Q una matriz tal que $QX_1 = I$ y $C = DQV_1$.

Supóngase que para este modelo σ_1^2 y σ_2^2 son desconocidos, pero que la razón $\rho = \sigma_2^2/\sigma_1^2$ es conocida, entonces se obtiene el modelo:

$$\begin{bmatrix} y_1 \\ y_2^* \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1^* \\ \varepsilon_2^* \end{bmatrix} \quad (4.16)$$

$$y^* = X^*\beta + \varepsilon^*$$

con

$$Cov(\varepsilon^*) = V_{\varepsilon^*} = \sigma_1^2 V_\rho = \sigma_1^2 \begin{bmatrix} V_1 & C' \\ -C & CV_1^{-1}C' + \rho V_2 \end{bmatrix}$$

donde $\rho \in \varrho$ y ϱ es el conjunto de los números reales positivos.

De esta manera, el modelo (4.16) es de la forma de un modelo de Gaus-Markov y según Bhimasankaram y Sengupta (1996), el BLUE de $X^*\beta$ del modelo $(y^*, X^*\beta, V_{\varepsilon^*} = \sigma_1^2 V_\rho)$ es obtenido como

$$\widehat{X^*\beta} = [I - V_\rho(I - P_{X^*})]\{(I - P_{X^*})V_\rho(I - P_{X^*})\}^-(I - P_{X^*})y^* \quad (4.17)$$

donde

$$P_{X^*} = \begin{bmatrix} P_{X_1} & 0 \\ 0 & P_{X_2} \end{bmatrix},$$

y P_{X_1} , P_{X_2} son las matrices de operador de proyección perpendicular sobre los subespacios $\mathcal{M}(X_1)$, $\mathcal{M}(X_2)$ respectivamente, dadas por $P_{X_i} = X_i(X_i'X_i)^{-1}X_i'$ para $i = 1, 2$.

Sea

$$Z = I - P_{X^*} = \begin{bmatrix} I_1 - P_{X_1} & 0 \\ 0 & I_2 - P_{X_2} \end{bmatrix} = \begin{bmatrix} Z_1 & 0 \\ 0 & Z_2 \end{bmatrix}$$

entonces, $X_1'Z_1 = 0$ y $X_2'Z_2 = 0$ y en consecuencia

$$i) V_\rho(I - P_{X^*}) = \begin{bmatrix} V_1 & -C' \\ -C & CV_1^{-1}C' + \rho V_2 \end{bmatrix} \begin{bmatrix} Z_1 & 0 \\ 0 & Z_2 \end{bmatrix}$$

$$\begin{aligned}
&= \begin{bmatrix} V_1 Z_1 & -C' Z_2 \\ -C Z_1 & C V_1^{-1} C' Z_2 + \rho V_2 Z_2 \end{bmatrix} \\
\text{ii) } (I - P_{X^*}) V_\rho (I - P_{X^*}) &= \begin{bmatrix} Z_1 & 0 \\ 0 & Z_2 \end{bmatrix} \begin{bmatrix} V_1 Z_1 & -C' Z_2 \\ -C Z_1 & C V_1^{-1} C' Z_2 + \rho V_2 Z_2 \end{bmatrix} \\
&= \begin{bmatrix} Z_1 V_1 Z_1 & -Z_1 C' Z_2 \\ -Z_2 C Z_1 & Z_2 C V_1^{-1} C' Z_2 + \rho Z_2 V_2 Z_2 \end{bmatrix}
\end{aligned}$$

La condición necesaria y suficiente para la existencia de los UBLUE dada por $\mathcal{M}(D) \subset \mathcal{M}(X_2)$, implica que $D' Z_2 = 0$ por ser $\mathcal{M}(X_2) \perp \mathcal{M}(Z_2)$ o equivalentemente $Z_2 D = 0$ y en consecuencia $Z_2 C = Z_2 D Q V_1 = 0$. De esta manera, se obtiene:

$$\begin{aligned}
\text{i) } V_\rho (I - P_{X^*}) &= \begin{bmatrix} V_1 Z_1 & 0 \\ -C Z_1 & \rho V_2 Z_2 \end{bmatrix} \\
\text{ii) } (I - P_{X^*}) V_\rho (I - P_{X^*}) &= \begin{bmatrix} Z_1 V_1 Z_1 & 0 \\ 0 & \rho Z_2 V_2 Z_2 \end{bmatrix} \\
\text{(iii) } \{(I - P_{X^*}) V_\rho (I - P_{X^*})\}^- &= \begin{bmatrix} \{Z_1 V_1 Z_1\}^- & 0 \\ 0 & (1/\rho) \{Z_2 V_2 Z_2\}^- \end{bmatrix} \\
\text{(iv) } V_\rho (I - P_{X^*}) \{(I - P_{X^*}) V_\rho (I - P_{X^*})\}^- (I - P_{X^*}) &= \\
&= \begin{bmatrix} V_1 Z_1 & 0 \\ -C Z_1 & \rho V_2 Z_2 \end{bmatrix} \begin{bmatrix} \{Z_1 V_1 Z_1\}^- & 0 \\ 0 & (1/\rho) \{Z_2 V_2 Z_2\}^- \end{bmatrix} \begin{bmatrix} Z_1 & 0 \\ 0 & Z_2 \end{bmatrix} \\
&= \begin{bmatrix} V_1 Z_1 \{Z_1 V_1 Z_1\}^- & 0 \\ -C Z_1 \{Z_1 V_1 Z_1\}^- & V_2 Z_2 \{Z_2 V_2 Z_2\}^- \end{bmatrix} \begin{bmatrix} Z_1 & 0 \\ 0 & Z_2 \end{bmatrix} \\
&= \begin{bmatrix} V_1 Z_1 \{Z_1 V_1 Z_1\}^- Z_1 & 0 \\ -D Q V_1 Z_1 \{Z_1 V_1 Z_1\}^- Z_1 & V_2 Z_2 \{Z_2 V_2 Z_2\}^- Z_2 \end{bmatrix}
\end{aligned}$$

En consecuencia, el BLUE de $X^*\beta$ esta dado por:

$$\begin{aligned}
 \widehat{X^*\beta} &= [I - V_\rho (I - P_{X^*}) \{(I - P_{X^*}) V_\rho (I - P_{X^*})\}^- (I - P_{X^*})] y^* \\
 &= \left\{ \begin{bmatrix} I_1 & 0 \\ 0 & I_2 \end{bmatrix} - \begin{bmatrix} V_1 Z_1 \{Z_1 V_1 Z_1\}^- Z_1 & 0 \\ -DQV_1 Z_1 \{Z_1 V_1 Z_1\}^- Z_1 & V_2 Z_2 \{Z_2 V_2 Z_2\}^- Z_2 \end{bmatrix} \right\} \begin{bmatrix} y_1 \\ y_2^* \end{bmatrix} \\
 &= \begin{bmatrix} I_1 - V_1 Z_1 \{Z_1 V_1 Z_1\}^- Z_1 & 0 \\ DQV_1 Z_1 \{Z_1 V_1 Z_1\}^- Z_1 & I_2 - V_2 Z_2 \{Z_2 V_2 Z_2\}^- Z_2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2^* \end{bmatrix} \\
 &= \begin{bmatrix} I_1 - P'_{Z_1 V_1} & 0 \\ DQP'_{Z_1 V_1} & I_2 - P'_{Z_2 V_2} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2^* \end{bmatrix}
 \end{aligned}$$

el cual es también el UBLUE para $X^*\beta$ en el modelo (4.16) como consecuencia de la condición necesaria y suficiente establecida por Volaufova (1987).

Por un procedimiento similar al empleado en la ecuación (4.2), durante la demostración del Lema 25, o bien de las partes (a) y (b) del Teorema 29, se tiene que:

- $[I_i - P'_{Z_i V_i}] V_i Z_i = P_{X_i | V_i Z_i} V_i Z_i = 0$
- $P'_{Z_i V_i} V_i Z_i = V_i Z_i$
- $[I_i - P'_{Z_i V_i}] X_i = P_{X_i | V_i Z_i} X_i = X_i$
- $P'_{Z_i V_i} X_i = 0$

para $i = 1, 2$. Luego se cumple:

$$\begin{aligned}
 DQP'_{Z_1 V_1} V_1 Z_1 - [I_2 - P'_{Z_2 V_2}] DQV_1 Z_1 &= DQV_1 Z_1 - DQV_1 Z_1 + P'_{Z_2 V_2} DQV_1 Z_1 \\
 &= P'_{Z_2 V_2} DQV_1 Z_1 = 0
 \end{aligned}$$

ya que $P'_{Z_2 V_2} D = 0$ por ser $Z_2 D = 0$ y por lo tanto se satisface:

$$\begin{aligned}
 \text{a) } &\begin{bmatrix} I_1 - P'_{Z_1 V_1} & 0 \\ DQP'_{Z_1 V_1} & I_2 - P'_{Z_2 V_2} \end{bmatrix} \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \\
 \text{b) } &\begin{bmatrix} I_1 - P'_{Z_1 V_1} & 0 \\ DQP'_{Z_1 V_1} & I_2 - P'_{Z_2 V_2} \end{bmatrix} \begin{bmatrix} V_1 & -C' \\ -C & CV_1^{-1}C' + \rho V_2 \end{bmatrix} \begin{bmatrix} Z_1 & 0 \\ 0 & Z_2 \end{bmatrix}
 \end{aligned}$$

$$\begin{aligned}
&= \begin{bmatrix} I_1 - P'_{Z_1 V_1} & 0 \\ DQP'_{Z_1 V_1} & I_2 - P'_{Z_2 V_2} \end{bmatrix} \begin{bmatrix} V_1 Z_1 & -C' Z_2 \\ -C Z_1 & CV_1^{-1} C' Z_2 + \rho V_2 Z_2 \end{bmatrix} \\
&= \begin{bmatrix} I_1 - P'_{Z_1 V_1} & 0 \\ DQP'_{Z_1 V_1} & I_2 - P'_{Z_2 V_2} \end{bmatrix} \begin{bmatrix} V_1 Z_1 & 0 \\ -DQV_1 Z_1 & \rho V_2 Z_2 \end{bmatrix} \\
&= \begin{bmatrix} [I_1 - P'_{Z_1 V_1}] V_1 Z_1 & 0 \\ DQP'_{Z_1 V_1} V_1 Z_1 - [I_2 - P'_{Z_2 V_2}] DQV_1 Z_1 & \rho [I_2 - P'_{Z_2 V_2}] V_2 Z_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}
\end{aligned}$$

En consecuencia, se denotará con $P_{X^*|V_\rho Z}$ a la matriz

$$\begin{bmatrix} I_1 - P'_{Z_1 V_1} & 0 \\ DQP'_{Z_1 V_1} & I_2 - P'_{Z_2 V_2} \end{bmatrix}$$

por ser un operador de proyección sobre $\mathcal{M}(X^*)$ a lo largo de $\mathcal{M}(V_\rho Z)$.

De esta manera, el estimador $\widehat{X^* \beta}$ se escribe como:

$$\widehat{X^* \beta} = P_{X^*|V_\rho Z} y^* \quad (4.18)$$

Obsérvese que $\widehat{X^* \beta}$ existe para todos los posibles valores de ρ , y por consiguiente de acuerdo con Wulff y Birkes (2005), se verifica que

$$\widehat{X^* \beta} = P_{X^*|V_\rho Z} y^* \quad (4.19)$$

es también un estimador UBLUE para $X^* \beta$.

Comentario 31 Obsérvese que usando el Lema (2.1) de Bhimasankaram and Sengupta (1996), la expresión para el BLUE de $X_1 \beta_1$ obtenida a partir del modelo $y_1 = X_1 \beta_1 + \varepsilon_1$, donde X_1 es de rango completo por columnas, $Cov(\varepsilon_1) = \sigma_1^2 V_1$ con V_1 no singular es

$$\widetilde{X_1 \beta_1} = (I_1 - P'_{Z_1 V_1}) y_1.$$

De igual forma, el BLUE de $\widetilde{X_2 \beta_2}$ obtenido a partir del modelo $y_2^* = X_2 \beta_2 + \varepsilon_2^*$, $V_{\varepsilon_2^*} = Cov(\varepsilon_2^*) = Cov(\varepsilon_2 - DQ\varepsilon_1) = Cov(\varepsilon_2) + Cov(DQ\varepsilon_1) = \sigma_2^2 V_2 + \sigma_1^2 DQV_1 Q' D'$, donde

X_2 es de rango completo por columnas, y con V_2 no singular es

$$\begin{aligned}\widetilde{X}_2\beta_2 &= \left[I_2 - V_{\varepsilon_2}^* (I_2 - P_{X_2}) \{ (I_2 - P_{X_2}) V_{\varepsilon_2}^* (I_2 - P_{X_2}) \}^- (I_2 - P_{X_2}) \right] y_2^* \\ \widetilde{X}_2\beta_2 &= (I_2 - P'_{Z_2V_2}) y_2^*\end{aligned}$$

4.4. Los UBLUE de β y $X\beta$ cuando $\rho = \sigma_2^2/\sigma_1^2$ es conocido

Ya que X_1 y X_2 son matrices de rango completo por columnas, de (4.19) se sigue que los UBLUE para β y $X\beta$ del modelo (4.13) están dados por:

$$\widehat{\beta} = X^{*-} \widehat{X^*}\beta = X^{*-} P_{X^*|V_\rho Z} y^* = X^{*-} P_{X^*|V_\rho Z} Fy \quad (4.20)$$

$$\widehat{X\beta} = X X^{*-} \widehat{X^*}\beta = X X^{*-} P_{X^*|V_\rho Z} y^* = X X^{*-} P_{X^*|V_\rho Z} Fy \quad (4.21)$$

donde

$$X^{*-} = \begin{bmatrix} X_1^- & 0 \\ 0 & X_2^- \end{bmatrix} = \begin{bmatrix} (X_1' X_1)^{-1} X_1' & 0' \\ 0 & (X_2' X_2)^{-1} X_2' \end{bmatrix}$$

En consecuencia, se obtiene

$$\begin{aligned}\widehat{\beta} &= \begin{bmatrix} X_1^- & 0 \\ 0 & X_2^- \end{bmatrix} \begin{bmatrix} I_1 - P'_{Z_1V_1} & 0 \\ -DQ(I_1 - P'_{Z_1V_1}) & I_2 - P'_{Z_2V_2} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \\ \widehat{X\beta} &= \begin{bmatrix} P_{X_1} & 0 \\ DX_1^- & P_{X_2} \end{bmatrix} \begin{bmatrix} I_1 - P'_{Z_1V_1} & 0 \\ -DQ(I_1 - P'_{Z_1V_1}) & I_2 - P'_{Z_2V_2} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}\end{aligned}$$

son los estimadores de β y $X\beta$ para el modelo en dos etapas regular bajo el supuesto que la razón $\rho = \sigma_2^2/\sigma_1^2$ es conocida. Obsérvese también que ambos estimadores son función de Fy .

Una expresión equivalente a (4.21) para el estimador de $\widehat{X\beta}$ esta dada por

$$\widehat{X\beta} = F^{-1} P_{X^*|V_\rho Z} Fy$$

ya que $XX^{*-} = F^{-1}X^*X^{*-}$ y $X^*X^{*-}P_{X^*|V_\rho Z} = P_{X^*|V_\rho Z}$.

Dado que $Z = F'Z$ también se cumple

$$a) \quad XX^{*-}P_{X^*|V_\rho Z}FX = XX^{*-}P_{X^*|V_\rho Z}X^* = XX^{*-}X^* = X.$$

$$b) \quad XX^{*-}P_{X^*|V_\rho Z}FV_\epsilon Z = XX^{*-}P_{X^*|V_\rho Z}FV_\epsilon F'Z = XX^{*-}P_{X^*|V_\rho Z}V_{\epsilon^*}Z = \\ \sigma_1^2 XX^{*-}P_{X^*|V_\rho Z}V_\rho Z = \sigma_1^2 XX^{*-}0 = 0.$$

Usando a) y b) se aprecia que $XX^{*-}P_{X^*|V_\rho Z}F$ (o equivalentemente $F^{-1}P_{X^*|V_\rho Z}F$) es un operador de proyección sobre $\mathcal{M}(X)$ paralelo (o a lo largo de) $\mathcal{M}(V_\epsilon Z)$.

También los estimadores (4.20) y (4.21) son UBLUE para β y $X\beta$ respectivamente para el modelo (4.13), al ser función del estimador dados en (4.19) cuando $\rho = \sigma_2^2/\sigma_1^2$ es conocido.

4.5. Los UBLUE de β y $X\beta$ para todo $\sigma_1^2 > 0, \sigma_2^2 > 0$

El UBLUE para $X\beta$ en el modelo (4.13) para todo $\sigma_1^2 > 0, \sigma_2^2 > 0$ puede ser obtenido a partir del UBLUE de $X^*\beta$ en el modelo (4.14) utilizando V_{ϵ^*} en (4.17) y siguiendo un procedimiento similar al aplicado para obtener el UBLUE de $X\beta$ a partir del modelo (4.16). Este y otros resultados se dan formalmente en el teorema a continuación.

Teorema 32 *Considérese el modelo lineal regular en dos etapas dado en (4.13). Si la condición $\mathcal{M}(D) \subset \mathcal{M}(X_2)$ se satisface, entonces*

1. *El proyector sobre $\mathcal{M}(X^*)$ paralelo a (o a lo largo de) $\mathcal{M}(V_\rho Z)$ es*

$$P_{X^*|V_\rho Z} = \begin{bmatrix} I_1 - P'_{Z_1 V_1} & 0 \\ DQP'_{Z_1 V_1} & I_2 - P'_{Z_2 V_2} \end{bmatrix}$$

2. *El BLUE de $X^*\beta$ obtenido a partir del modelo (4.16) está dado por*

$$\widehat{X^*\beta} = R_{X^*|V_\rho Z}y^*.$$

3. *El BLUE $\widehat{X^*\beta} = P_{X^*|V_\rho Z}y^*$ es también el UBLUE de $X^*\beta$.*

4. Los UBLUE de β y $X\beta$ para el modelo (4.13) cuando $\rho = \sigma_2^2/\sigma_1^2$ es conocido, son respectivamente

$$\begin{aligned}\widehat{\beta} &= X^{*-} R_{X^*|V_\rho Z} F y \\ \widehat{X\beta} &= X X^{*-} R_{X^*|V_\rho Z} F y\end{aligned}$$

5. Los UBLUE de β y $X\beta$ para el modelo (4.13) para todo $\sigma_1^2 > 0$, $\sigma_2^2 > 0$ son

$$\begin{aligned}\widehat{\beta} &= X^{*-} R_{X^*|V_\epsilon Z} F y \\ \widehat{X\beta} &= X X^{*-} R_{X^*|V_\epsilon Z} F y\end{aligned}$$

los cuales son iguales a los estimadores correspondientes al caso cuando $\rho = \sigma_2^2/\sigma_1^2$ es conocido.

6. El proyector sobre $\mathcal{M}(X)$ paralelo a (o a lo largo de) $\mathcal{M}(V_\epsilon Z)$ es

$$X X^{*-} P_{X^*|V_\epsilon Z} F.$$

Demostración.

1. La demostración que

$$P_{X^*|V_\rho Z} = \begin{bmatrix} I_1 - P'_{Z_1 V_1} & 0 \\ DQ P'_{Z_1 V_1} & I_2 - P'_{Z_2 V_2} \end{bmatrix}$$

es el operador de proyección sobre $\mathcal{M}(X^*)$ paralelo a (o a lo largo de) $\mathcal{M}(V_\rho Z)$ es consecuencia de la verificación de las condiciones necesarias y suficientes dadas por Rao (1974) en el Lema 2.4 de su artículo, según las cuales, para la partición disjunta $(X^* : V_\rho Z)$, la matriz $P_{X^*|V_\rho Z}$ es un proyector sobre $\mathcal{M}(X^*)$ paralelo a (o a lo largo de) $\mathcal{M}(V_\rho Z)$ si y solo si las condiciones

- a) $P_{X^*|V_\rho Z} X^* = X^*$

$$\text{b) } P_{X^*|V_\rho Z} V_\rho Z = 0$$

se satisfacen, las cuales se aprecia que se verifican a través de los productos de matrices a continuación:

$$\begin{aligned} \text{a) } & \begin{bmatrix} I_1 - P'_{Z_1 V_1} & 0 \\ DQP'_{Z_1 V_1} & I_2 - P'_{Z_2 V_2} \end{bmatrix} \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \\ \text{b) } & \begin{bmatrix} I_1 - P'_{Z_1 V_1} & 0 \\ DQP'_{Z_1 V_1} & I_2 - P'_{Z_2 V_2} \end{bmatrix} \begin{bmatrix} V_1 & -C' \\ -C & CV_1^{-1}C' + \rho V_2 \end{bmatrix} \begin{bmatrix} Z_1 & 0 \\ 0 & Z_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \end{aligned}$$

2. Este resultado es obtenido por la aplicación de la expresión (4.7) al modelo (4.16) para obtener el BLUE de $X^*\beta$ dado en (4.17). Específicamente, para obtener

$$\widehat{X^*\beta} = [I - V_\rho(I - P_{X^*})]\{(I - P_{X^*})V_\rho(I - P_{X^*})\}^-(I - P_{X^*})y^*$$

Luego siguiendo el procedimiento desarrollado en las páginas 54 a 57 de este trabajo, bajo la condición $\mathcal{M}(D) \subset \mathcal{M}(X_2)$ se obtiene la expresión (4.18) del BLUE de $X^*\beta$ dada por:

$$\widehat{X^*\beta} = R_{X^*|V_\rho Z} y^*.$$

3. Ya que el BLUE $\widehat{X^*\beta} = P_{X^*|V_\rho Z} y^*$ de $X^*\beta$ obtenido en la parte (2) de este teorema existe para todo $\rho \in \varrho$, de acuerdo a Wulff y Birkes (2005) es también el *UBBLUE* de $X^*\beta$.

Otra manera de probar este resultado es considerando que de acuerdo al Teorema 1.2 de Volaufova (1987), el *UBLE* del vector $\beta = (\beta'_1, \beta'_2)'$ del modelo (4.16) existe si y solo si $\mathcal{M}(D) \subset \mathcal{M}(X_2)$, la cual es la condición utilizada junto con la ecuación (4.17) para obtener el BLUE de $X^*\beta$ del modelo (4.16) en la parte (2) de este teorema y dado que $E(X^{*-}\widehat{X^*\beta}) = X^{*-}E(\widehat{X^*\beta}) = X^{*-}X^*\beta = \beta$, se sigue que $\widehat{\beta} = X^{*-}\widehat{X^*\beta}$ y $\widehat{X^*\beta} = X^*\widehat{\beta}$. Por consiguiente $\widehat{\beta}$ y $\widehat{X^*\beta}$ están relacionados linealmente y puesto que $\widehat{\beta} = X^{*-}\widehat{X^*\beta}$ es el *UBBLUE* de β para el modelo (4.16), se sigue que $\widehat{X^*\beta}$ es también el *UBBLUE* para $X^*\beta$.

4. Dado que $\widehat{X^*\beta}$ existe para todos los posibles valores de ρ , según Wulff y Birkes

(2005) el estimador

$$\widehat{X^*\beta} = P_{X^*|V_\rho Z} y^*$$

es también un estimador UBLUE para $X^*\beta$. De los items a) y b) de la página 59 se sigue que al pre-multiplicar esta ecuación por X^{*-} y XX^{*-} en cada caso, se obtienen los estimadores

$$\begin{aligned}\widehat{\beta} &= X^{*-} \widehat{X^*\beta} = X^{*-} P_{X^*|V_\rho Z} y^* = X^{*-} P_{X^*|V_\rho Z} Fy \\ \widehat{X\beta} &= XX^{*-} \widehat{X^*\beta} = XX^{*-} P_{X^*|V_\rho Z} y^* = XX^{*-} P_{X^*|V_\rho Z} Fy\end{aligned}$$

los cuales son los UBLUE de β y $X\beta$ para el modelo (4.13) para el caso cuando $\rho = \sigma_2^2/\sigma_1^2$ es conocido.

5. El resultado se sigue de la aplicación la expresión (4.7) al modelo (4.15) y así obtener el BLUE de $X^*\beta$ dado por

$$\widehat{X^*\beta} = [I - V_{\epsilon^*}(I - P_{X^*})]\{(I - P_{X^*})V_{\epsilon^*}(I - P_{X^*})\}^{-1}(I - P_{X^*})y^*$$

luego, considerando la condición $\mathcal{M}(D) \subset \mathcal{M}(X_2)$ y siguiendo un procedimiento igual al llevado a cabo en las páginas 54 a 57 para obtener el BLUE de $X^*\beta$ para el modelo (4.16) se obtiene:

$$\begin{aligned}\text{i) } V_{\epsilon^*}(I - P_{X^*}) &= \begin{bmatrix} \sigma_1^2 V_1 & -\sigma_1^2 C' \\ -\sigma_1^2 C & \sigma_1^2 C V_1^{-1} C' + \sigma_2^2 V_2 \end{bmatrix} \begin{bmatrix} Z_1 & 0 \\ 0 & Z_2 \end{bmatrix} \\ &= \begin{bmatrix} \sigma_1^2 V_1 Z_1 & -\sigma_1^2 C' Z_2 \\ -\sigma_1^2 C Z_1 & \sigma_1^2 C V_1^{-1} C' Z_2 + \sigma_2^2 V_2 Z_2 \end{bmatrix} \\ \text{ii) } (I - P_{X^*}) V_{\epsilon^*} (I - P_{X^*}) &= \begin{bmatrix} Z_1 & 0 \\ 0 & Z_2 \end{bmatrix} \begin{bmatrix} \sigma_1^2 V_1 Z_1 & -\sigma_1^2 C' Z_2 \\ -\sigma_1^2 C Z_1 & \sigma_1^2 C V_1^{-1} C' Z_2 + \sigma_2^2 V_2 Z_2 \end{bmatrix} \\ &= \begin{bmatrix} \sigma_1^2 Z_1 V_1 Z_1 & -\sigma_1^2 Z_1 C' Z_2 \\ -\sigma_1^2 Z_2 C Z_1 & \sigma_1^2 Z_2 C V_1^{-1} C' Z_2 + \sigma_2^2 Z_2 V_2 Z_2 \end{bmatrix}\end{aligned}$$

De la condición $\mathcal{M}(D) \subset \mathcal{M}(X_2)$, se tiene que $D'Z_2 = 0$ al ser $\mathcal{M}(X_2) \perp \mathcal{M}(Z_2)$ o equivalentemente $Z_2 D = 0$ y por lo tanto $Z_2 C = Z_2 D Q V_1 = 0$. De esta manera,

se obtiene:

$$\begin{aligned}
 \text{i) } V_{\varepsilon^*} (I - P_{X^*}) &= \begin{bmatrix} \sigma_1^2 V_1 Z_1 & 0 \\ -\sigma_1^2 C Z_1 & \sigma_2^2 V_2 Z_2 \end{bmatrix} \\
 \text{ii) } (I - P_{X^*}) V_{\varepsilon^*} (I - P_{X^*}) &= \begin{bmatrix} \sigma_1^2 Z_1 V_1 Z_1 & 0 \\ 0 & \sigma_2^2 Z_2 V_2 Z_2 \end{bmatrix}. \\
 \text{iii) } \{(I - P_{X^*}) V_{\varepsilon^*} (I - P_{X^*})\}^- &= \begin{bmatrix} (1/\sigma_1^2) \{Z_1 V_1 Z_1\}^- & 0 \\ 0 & (1/\sigma_2^2) \{Z_2 V_2 Z_2\}^- \end{bmatrix} \\
 \text{(iv) } V_{\varepsilon^*} (I - P_{X^*}) \{(I - P_{X^*}) V_{\varepsilon^*} (I - P_{X^*})\}^- (I - P_{X^*}) &= \\
 &= \begin{bmatrix} \sigma_1^2 V_1 Z_1 & 0 \\ -\sigma_1^2 C Z_1 & \sigma_2^2 V_2 Z_2 \end{bmatrix} \begin{bmatrix} \{Z_1 V_1 Z_1\}^- / \sigma_1^2 & 0 \\ 0 & \{Z_2 V_2 Z_2\}^- / \sigma_2^2 \end{bmatrix} \begin{bmatrix} Z_1 & 0 \\ 0 & Z_2 \end{bmatrix} \\
 &= \begin{bmatrix} V_1 Z_1 \{Z_1 V_1 Z_1\}^- & 0 \\ -C Z_1 \{Z_1 V_1 Z_1\}^- & V_2 Z_2 \{Z_2 V_2 Z_2\}^- \end{bmatrix} \begin{bmatrix} Z_1 & 0 \\ 0 & Z_2 \end{bmatrix} \\
 &= \begin{bmatrix} V_1 Z_1 \{Z_1 V_1 Z_1\}^- Z_1 & 0 \\ -DQV_1 Z_1 \{Z_1 V_1 Z_1\}^- Z_1 & V_2 Z_2 \{Z_2 V_2 Z_2\}^- Z_2 \end{bmatrix}
 \end{aligned}$$

De esta manera, el BLUE de $X^* \beta$ para el modelo (4.15) está dado por:

$$\begin{aligned}
 \widehat{X^* \beta} &= [I - V_{\varepsilon^*} (I - P_{X^*}) \{(I - P_{X^*}) V_{\varepsilon^*} (I - P_{X^*})\}^- (I - P_{X^*})] y^* \\
 &= \left\{ \begin{bmatrix} I_1 & 0 \\ 0 & I_2 \end{bmatrix} - \begin{bmatrix} V_1 Z_1 \{Z_1 V_1 Z_1\}^- Z_1 & 0 \\ -DQV_1 Z_1 \{Z_1 V_1 Z_1\}^- Z_1 & V_2 Z_2 \{Z_2 V_2 Z_2\}^- Z_2 \end{bmatrix} \right\} \begin{bmatrix} y_1 \\ y_2^* \end{bmatrix} \\
 &= \begin{bmatrix} I_1 - V_1 Z_1 \{Z_1 V_1 Z_1\}^- Z_1 & 0 \\ DQV_1 Z_1 \{Z_1 V_1 Z_1\}^- Z_1 & I_2 - V_2 Z_2 \{Z_2 V_2 Z_2\}^- Z_2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2^* \end{bmatrix} \\
 &= \begin{bmatrix} I_1 - P'_{Z_1 V_1} & 0 \\ DQP'_{Z_1 V_1} & I_2 - P'_{Z_2 V_2} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2^* \end{bmatrix} \\
 &= R_{X^*|V_{\varepsilon^*} Z} y^*
 \end{aligned}$$

Por simple comparación es fácil verificar que $P_{X^*|V_{\rho} Z} = P_{X^*|V_{\varepsilon^*} Z}$. Por lo tanto el estimador de $X^* \beta$ cuando $\sigma_1^2 > 0$, $\sigma_2^2 > 0$ es igual al estimador para el caso cuando $\rho = \sigma_2^2 / \sigma_1^2$ es conocido, y dado que $\widehat{X^* \beta}$ existe para todos los posibles

valores de los parámetros σ_1^2, σ_2^2 , según Wulff y Birkes (2005) es también UBLUE. En consecuencia, los estimadores de β y $X\beta$ en los casos cuando $\sigma_1^2 > 0, \sigma_2^2 > 0$ y cuando $\rho = \sigma_2^2/\sigma_1^2$ conocido son iguales, son UBLUE y podemos escribirlos como:

$$\hat{\beta} = X^{*-} P_{X^*|V_{\epsilon^*}Z} Fy$$

$$\widehat{X\beta} = XX^{*-} P_{X^*|V_{\epsilon^*}Z} Fy$$

6. En las partes a) y b) de la página 59 se encontró que

$$XX^{*-} P_{X^*|V_{\rho}Z} FX = X$$

$$XX^{*-} P_{X^*|V_{\rho}Z} FV_{\epsilon}Z = 0$$

y de la prueba de la parte 5 de este teorema se sabe que $P_{X^*|V_{\rho}Z} = P_{X^*|V_{\epsilon^*}Z}$. Por lo tanto se obtiene que también

$$XX^{*-} P_{X^*|V_{\epsilon}Z} FX = X$$

$$XX^{*-} P_{X^*|V_{\epsilon}Z} FV_{\epsilon}Z = 0$$

y en consecuencia $XX^{*-} P_{X^*|V_{\epsilon}Z} F$ es un proyector oblicuo sobre $\mathcal{M}(X)$ paralelo a (o a lo largo de) $\mathcal{M}(V_{\epsilon}Z)$.

Obsérvese que el modelo (4.15) es una transformación no singular del modelo (4.13) y que la expresión para $\widehat{X\beta}$ es una función lineal de Fy y dado que éste es un BLUE (y también UBLUE), se sigue de la definición 3.1 dada por Drygas (1983) que Fy es una transformación lineal suficiente.

Comentario 33 *Los resultados obtenidos hasta ahora en este trabajo también pueden derivarse usando la expresión dada en el Teorema 3.2 parte (a) de Rao (1974). En este caso las matrices Z_1, Z_2 son matrices de máximo rango que satisfacen $Z_1 = X_1^{\perp}, Z_2 = X_2^{\perp}$.*

Capítulo 5

Generalización al modelo de p-etapas

En este capítulo se generalizan los resultados obtenidos para el modelo en 2-etapas al modelo en p-etapas. Para alcanzar este objetivo se introduce una metodología sencilla que permite encontrar la matriz de transformación no singular F del modelo lineal de p etapas regular y se demuestra que las expresiones para obtener el estimador UBLUE del vector de parámetros β y de $X\beta$ son similares a las encontradas para el caso del modelo lineal en dos etapas regular.

5.1. Relación entre el modelo lineal regular y el modelo transformado

5.1.1. Caso: Modelo en dos etapas

Sea el modelo lineal en dos etapas regular

$$y_1 = X_1\beta_1 + \varepsilon_1; \quad (5.1)$$

$$y_2 = C_{21}\beta_1 + X_2\beta_2 + \varepsilon_2$$

que matricialmente escribimos como

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ C_{21} & X_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}$$

y de manera compacta como $y = X\beta + \varepsilon$.

Sea Q_1 una matriz tal que $Q_1 X_1 = I_{p_1}$. Entonces $C_{21}\beta_1$ puede escribirse como:

$$C_{21}\beta_1 = C_{21}Q_1 X_1 \beta_1 = C_{21}Q_1 (y_1 - \varepsilon_1) = C_{21}Q_1 y_1 - C_{21}Q_1 \varepsilon_1$$

que al sustituir en el modelo (5.1) resulta:

$$\begin{aligned} y_1 &= X_1 \beta_1 + \varepsilon_1 \\ y_2 &= C_{21}Q_1 y_1 - C_{21}Q_1 \varepsilon_1 + X_2 \beta_2 + \varepsilon_2 \end{aligned}$$

o equivalentemente

$$\begin{aligned} y_1 &= X_1 \beta_1 + \varepsilon_1 \\ y_2 - C_{21}Q_1 y_1 &= X_2 \beta_2 - C_{21}Q_1 \varepsilon_1 + \varepsilon_2 \end{aligned} \quad (5.2)$$

En forma matricial al modelo (5.2) se escribe como:

$$\begin{bmatrix} I_1 & 0 \\ -C_{21}Q_1 & I_2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} I_1 & 0 \\ -C_{21}Q_1 & I_2 \end{bmatrix} \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}$$

$$\begin{bmatrix} I_1 & 0 \\ -F_{21} & I_2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} I_1 & 0 \\ -F_{21} & I_2 \end{bmatrix} \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix} \quad (5.3)$$

donde $F_{21} = C_{21}Q_1$.

Denominando

$$F = \begin{bmatrix} I_1 & 0 \\ -F_{21} & I_2 \end{bmatrix}$$

el modelo (5.3) puede expresarse como $Fy = FX\beta + F\varepsilon$ dado que $FX = X^*$, esto es:

$$\begin{bmatrix} I_1 & 0 \\ -C_{21}Q_1 & I_2 \end{bmatrix} \begin{bmatrix} X_1 & 0 \\ C_{21} & X_2 \end{bmatrix} = \begin{bmatrix} I_1 & 0 \\ -F_{21} & I_2 \end{bmatrix} \begin{bmatrix} X_1 & 0 \\ C_{21} & X_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix}$$

Obsérvese que $Fy = FX\beta + F\varepsilon$ es el modelo transformado obtenido al pre-multiplicar el modelo original $y = X\beta + \varepsilon$ por la matriz de transformación F . Específicamente el modelo transformado es $y^* = X^*\beta + \varepsilon^*$ donde $y^* = Fy$, $X^* = FX$ y $\varepsilon^* = F\varepsilon$.

5.1.2. Caso: Modelo en tres etapas

Sea el modelo en tres etapas

$$\begin{aligned} y_1 &= X_1\beta_1 + \varepsilon_1; \\ y_2 &= C_{21}\beta_1 + X_2\beta_2 + \varepsilon_2 \\ y_3 &= C_{31}\beta_1 + C_{32}\beta_2 + X_3\beta_3 + \varepsilon_3 \end{aligned} \tag{5.4}$$

y sean Q_j matrices tales que $Q_j X_j = I_{p_j}$, $j = 1, 2$. Siguiendo un procedimiento similar al de la sección anterior, se transforman los coeficientes $C_{21}\beta_1$, $C_{31}\beta_1$, $C_{32}\beta_2$ para este modelo:

$$\begin{aligned} C_{21}\beta_1 &= C_{21}Q_1 X_1\beta_1 = C_{21}Q_1 (y_1 - \varepsilon_1) = C_{21}Q_1 y_1 - C_{21}Q_1 \varepsilon_1 \\ C_{31}\beta_1 &= C_{31}Q_1 X_1\beta_1 = C_{31}Q_1 (y_1 - \varepsilon_1) = C_{31}Q_1 y_1 - C_{31}Q_1 \varepsilon_1 \\ C_{32}\beta_2 &= C_{32}Q_2 X_2\beta_2 = C_{32}Q_2 (y_2 - \varepsilon_2 - C_{21}\beta_1) = C_{32}Q_2 (y_2 - \varepsilon_2 - C_{21}Q_1 y_1 + C_{21}Q_1 \varepsilon_1) \\ &= -C_{32}Q_2 C_{21}Q_1 y_1 + C_{32}Q_2 y_2 + C_{32}Q_2 C_{21}Q_1 \varepsilon_1 - C_{32}Q_2 \varepsilon_2 \end{aligned}$$

que al ser sustituidos en (5.4) producen como resultado:

$$\begin{aligned}
 y_1 &= X_1\beta_1 + \varepsilon_1; \\
 -C_{21}Q_1y_1 + y_2 &= X_2\beta_2 - C_{21}Q_1\varepsilon_1 + \varepsilon_2 \\
 -C_{31}Q_1y_1 + C_{32}Q_2C_{21}Q_1y_1 - C_{32}Q_2y_2 + y_3 &= X_3\beta_3 - C_{31}Q_1\varepsilon_1 \\
 &\quad + C_{32}Q_2C_{21}Q_1\varepsilon_1 - C_{32}Q_2\varepsilon_2 + \varepsilon_3
 \end{aligned}$$

Denominando $F_{31} = C_{31}Q_1 - C_{32}Q_2C_{21}Q_1$ y $F_{32} = C_{32}Q_2$, se obtienen las ecuaciones:

$$\begin{aligned}
 y_1 &= X_1\beta_1 + \varepsilon_1; \\
 -F_{21}y_1 + y_2 &= X_2\beta_2 - F_{21}\varepsilon_1 + \varepsilon_2 \\
 -F_{31}y_1 - F_{32}y_2 + y_3 &= X_3\beta_3 - F_{31}\varepsilon_1 - F_{32}\varepsilon_2 + \varepsilon_3
 \end{aligned}$$

que en forma matricial se expresan como:

$$\begin{bmatrix} I_1 & 0 & 0 \\ -F_{21} & I_2 & 0 \\ -F_{31} & -F_{32} & I_3 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} X_1 & 0 & 0 \\ 0 & X_2 & 0 \\ 0 & 0 & X_3 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \begin{bmatrix} I_1 & 0 & 0 \\ -F_{21} & I_2 & 0 \\ -F_{31} & -F_{32} & I_3 \end{bmatrix} \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix}$$

De esta manera, para el caso de tres etapas, la matriz F esta dada por:

$$F_3 = \begin{bmatrix} I_1 & \mathbf{0} & \mathbf{0} \\ -F_{21} & I_2 & \mathbf{0} \\ -F_{31} & -F_{32} & I_3 \end{bmatrix}$$

5.1.3. Caso: Modelo en cuatro etapas

En el caso del modelo lineal en cuatro etapas regular

$$\begin{aligned}
 y_1 &= X_1\beta_1 + \varepsilon_1; \\
 y_2 &= C_{21}\beta_1 + X_2\beta_2 + \varepsilon_2 \\
 y_3 &= C_{31}\beta_1 + C_{32}\beta_2 + X_3\beta_3 + \varepsilon_3 \\
 y_4 &= C_{41}\beta_1 + C_{42}\beta_2 + C_{43}\beta_3 + X_4\beta_4 + \varepsilon_4
 \end{aligned} \tag{5.5}$$

los coeficientes del modelo se escriben como:

$$\begin{aligned}
 C_{21}\beta_1 &= C_{21}Q_1y_1 - C_{21}Q_1\varepsilon_1 \\
 C_{31}\beta_1 &= C_{31}Q_1y_1 - C_{31}Q_1\varepsilon_1 \\
 C_{41}\beta_1 &= C_{41}Q_1y_1 - C_{41}Q_1\varepsilon_1 \\
 C_{32}\beta_2 &= -C_{32}Q_2C_{21}Q_1y_1 + C_{32}Q_2y_2 + C_{32}Q_2C_{21}Q_1\varepsilon_1 - C_{32}Q_2\varepsilon_2 \\
 C_{42}\beta_2 &= -C_{42}Q_2C_{21}Q_1y_1 + C_{42}Q_2y_2 + C_{42}Q_2C_{21}Q_1\varepsilon_1 - C_{42}Q_2\varepsilon_2 \\
 C_{43}\beta_3 &= -C_{43}Q_3C_{31}Q_1y_1 + C_{43}Q_3C_{32}Q_2C_{21}Q_1y_1 - C_{43}Q_3C_{32}Q_2y_2 + C_{43}Q_3y_3 \\
 &\quad + C_{43}Q_3C_{31}Q_1\varepsilon_1 - C_{43}Q_3C_{32}Q_2C_{21}Q_1\varepsilon_1 + C_{43}Q_3C_{32}Q_2\varepsilon_2 - C_{43}Q_3\varepsilon_3
 \end{aligned}$$

y al ser sustituidos en (5.5) lo transforman en:

$$\begin{aligned}
 y_1 &= X_1\beta_1 + \varepsilon_1; \\
 -F_{21}y_1 + y_2 &= X_2\beta_2 - F_{21}\varepsilon_1 + \varepsilon_2 \\
 -F_{31}y_1 - F_{32}y_2 + y_3 &= X_3\beta_3 - F_{31}\varepsilon_1 - F_{32}\varepsilon_2 + \varepsilon_3 \\
 -F_{41}y_1 - F_{42}y_2 - F_{43}y_3 + y_4 &= X_4\beta_4 - F_{41}\varepsilon_1 - F_{42}\varepsilon_2 - F_{43}\varepsilon_3 + \varepsilon_4
 \end{aligned}$$

que de manera matricial escribimos:

$$\begin{bmatrix} I_1 & 0 & 0 & 0 \\ -F_{21} & I_2 & 0 & 0 \\ -F_{31} & -F_{32} & I_3 & 0 \\ -F_{41} & -F_{42} & -F_{43} & I_4 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} X_1 & 0 & 0 & 0 \\ 0 & X_2 & 0 & 0 \\ 0 & 0 & X_3 & 0 \\ 0 & 0 & 0 & X_4 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix} + \begin{bmatrix} I_1 & 0 & 0 & 0 \\ -F_{21} & I_2 & 0 & 0 \\ -F_{31} & -F_{32} & I_3 & 0 \\ -F_{41} & -F_{42} & -F_{43} & I_4 \end{bmatrix} \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{bmatrix}$$

con los términos F_{21} , F_{31} , F_{41} , F_{32} , F_{42} y F_{43} dados por:

$$F_{21} = C_{21}Q_1$$

$$F_{31} = C_{31}Q_1 - C_{32}Q_2C_{21}Q_1$$

$$F_{41} = C_{41}Q_1 - C_{42}Q_2C_{21}Q_1 - C_{43}Q_3C_{31}Q_1 + C_{43}Q_3C_{32}Q_2C_{21}Q_1$$

$$F_{32} = C_{32}Q_2$$

$$F_{42} = C_{42}Q_2 - C_{43}Q_3C_{32}Q_2$$

$$F_{43} = C_{43}Q_3$$

En este caso la matriz F de transformación es

$$F_4 = \begin{bmatrix} I_1 & 0 & 0 & 0 \\ -F_{21} & I_2 & 0 & 0 \\ -F_{31} & -F_{32} & I_3 & 0 \\ -F_{41} & -F_{42} & -F_{43} & I_4 \end{bmatrix}$$

5.1.4. Caso: Modelo en cinco etapas

Cuando el modelo lineal posee 5-etapas, las ecuaciones que se obtienen son:

$$C_{21}\beta_1 = C_{21}Q_1y_1 - C_{21}Q_1\varepsilon_1$$

$$C_{31}\beta_1 = C_{31}Q_1y_1 - C_{31}Q_1\varepsilon_1$$

$$C_{41}\beta_1 = C_{41}Q_1y_1 - C_{41}Q_1\varepsilon_1$$

$$C_{51}\beta_1 = C_{51}Q_1y_1 - C_{51}Q_1\varepsilon_1$$

$$C_{32}\beta_2 = -C_{32}Q_2C_{21}Q_1y_1 + C_{32}Q_2y_2 + C_{32}Q_2C_{21}Q_1\varepsilon_1 - C_{32}Q_2\varepsilon_2$$

$$C_{42}\beta_2 = -C_{42}Q_2C_{21}Q_1y_1 + C_{42}Q_2y_2 + C_{42}Q_2C_{21}Q_1\varepsilon_1 - C_{42}Q_2\varepsilon_2$$

$$C_{52}\beta_2 = -C_{52}Q_2C_{21}Q_1y_1 + C_{52}Q_2y_2 + C_{52}Q_2C_{21}Q_1\varepsilon_1 - C_{52}Q_2\varepsilon_2$$

$$\begin{aligned} C_{43}\beta_3 = & -C_{43}Q_3C_{31}Q_1y_1 + C_{43}Q_3C_{32}Q_2C_{21}Q_1y_1 - C_{43}Q_3C_{32}Q_2y_2 + C_{43}Q_3y_3 \\ & + C_{43}Q_3C_{31}Q_1\varepsilon_1 - C_{43}Q_3C_{32}Q_2C_{21}Q_1\varepsilon_1 - C_{43}Q_3C_{32}Q_2\varepsilon_2 - C_{43}Q_3\varepsilon_3 \end{aligned}$$

$$\begin{aligned} C_{53}\beta_3 = & -C_{53}Q_3C_{31}Q_1y_1 + C_{53}Q_3C_{32}Q_2C_{21}Q_1y_1 - C_{53}Q_3C_{32}Q_2y_2 + C_{53}Q_3y_3 \\ & + C_{53}Q_3C_{31}Q_1\varepsilon_1 - C_{53}Q_3C_{32}Q_2C_{21}Q_1\varepsilon_1 + C_{53}Q_3C_{32}Q_2\varepsilon_2 - C_{53}Q_3\varepsilon_3 \end{aligned}$$

$$\begin{aligned} C_{54}\beta_4 = & -C_{54}Q_4C_{41}Q_1y_1 + C_{54}Q_4C_{42}Q_2C_{21}Q_1y_1 - C_{54}Q_4C_{43}Q_3C_{32}Q_2C_{21}Q_1y_1 \\ & + C_{54}Q_4C_{43}Q_3C_{31}Q_1y_1 - C_{54}Q_4C_{42}Q_2y_2 + C_{54}Q_4C_{43}Q_3C_{32}Q_2y_2 - C_{54}Q_4C_{43}Q_3y_3 \\ & + C_{54}Q_4y_4 + C_{54}Q_4C_{41}Q_1\varepsilon_1 - C_{54}Q_4C_{42}Q_2C_{21}Q_1\varepsilon_1 + C_{54}Q_4C_{43}Q_3C_{32}Q_2C_{21}Q_1\varepsilon_1 \\ & - C_{54}Q_4C_{43}Q_3C_{31}Q_1\varepsilon_1 + C_{54}Q_4C_{42}Q_2\varepsilon_2 - C_{54}Q_4C_{43}Q_3C_{32}Q_2\varepsilon_2 + C_{54}Q_4C_{43}Q_3\varepsilon_3 \\ & - C_{54}Q_4\varepsilon_4 \end{aligned}$$

y la matriz de transformación F del modelo lineal en 5-etapas regular en el modelo transformado es:

$$F_5 = \begin{bmatrix} I_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ -F_{21} & I_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ -F_{31} & -F_{32} & I_3 & \mathbf{0} & \mathbf{0} \\ -F_{41} & -F_{42} & -F_{43} & I_4 & \mathbf{0} \\ -F_{51} & -F_{52} & -F_{53} & -F_{54} & I_5 \end{bmatrix}$$

donde los términos F_{ij} para $i = 2, 3, 4, 5$ y $j = 1, 2, 3, 4$ e $i > j$ son:

$$F_{21} = C_{21}Q_1$$

$$F_{31} = C_{31}Q_1 - C_{32}Q_2C_{21}Q_1$$

$$F_{41} = C_{41}Q_1 - C_{42}Q_2C_{21}Q_1 - C_{43}Q_3C_{31}Q_1 + C_{43}Q_3C_{32}Q_2C_{21}Q_1$$

$$F_{51} = C_{51}Q_1 - C_{52}Q_2C_{21}Q_1 - C_{53}Q_3C_{31}Q_1 + C_{53}Q_3C_{32}Q_2C_{21}Q_1 - C_{54}Q_4C_{41}Q_1 \\ + C_{54}Q_4C_{42}Q_2C_{21}Q_1 + C_{54}Q_4C_{43}Q_3C_{31}Q_1 - C_{54}Q_4C_{43}Q_3C_{32}Q_2C_{21}Q_1$$

$$F_{32} = C_{32}Q_2$$

$$F_{42} = C_{42}Q_2 - C_{43}Q_3C_{32}Q_2$$

$$F_{52} = C_{52}Q_2 - C_{53}Q_3C_{32}Q_2 - C_{54}Q_4C_{42}Q_2 + C_{54}Q_4C_{43}Q_3C_{32}Q_2$$

$$F_{43} = C_{43}Q_3$$

$$F_{53} = C_{53}Q_3 - C_{54}Q_4C_{43}Q_3$$

$$F_{54} = C_{54}Q_4$$

En este punto es importante destacar los siguientes aspectos:

1. Los términos en que se expresan los coeficientes $C_{ij}\beta_j$ depende del parámetro β_j involucrado, $i = 2, 3, 4, 5$; $j = 1, 2, 3, 4$; $i > j$. Esto es, la forma de los términos $C_{ij}\beta_j$ es la misma para un mismo j .
2. La matriz F de la etapa i contiene a la matriz F de la etapa anterior $i-1$; $i = 2, 3, 4, 5$
3. Las submatrices de F relacionadas con la etapa i ; esto es, las F_{ij} , $i = 2, 3, 4, 5$; $j = 1, 2, \dots, i-1$, son fácil de obtener a partir de los términos F_{ik} , para $j < k < i-1$ de la misma etapa i previamente calculados y de las matrices C_{ij} de etapas anteriores.

Los aspectos 1 y 2 son fácil de comprender por simple inspección. En cuanto al tercer aspecto, en la siguiente sección se da una explicación detallada del mismo.

5.2. Procedimiento iterativo para obtener F

El proceso de construcción de la matriz F para un modelo en p etapas es iterativo y depende de los términos de la matriz F para un modelo con $p-1$ etapas. Esta dependencia es lo que hace que el procedimiento sea iterativo; pero ello no debe preocuparnos pues el proceso es sumamente sencillo y fácil de implementar en cualquier lenguaje de computación. En este proceso se parte del hecho que las matrices C_{ij} y Q_j son conocidas y que las matrices Q_j cumplen con la condición $Q_j X_j = I_{p_j}$, para $i = 2, 3, \dots, p$; $j = 1, 2, \dots, p-1$; $i > j$, donde I_{p_j} es una matriz identidad de $p_j \times p_j$.

En lo que sigue se denotar con F_i a la matriz de transformación de un modelo lineal en i etapas regular $y = X\beta + \varepsilon$ en el modelo transformado $Fy = FX\beta + F\varepsilon$. Claramente, la matriz para el modelo lineal en dos etapas esta dada por

$$F_2 = \begin{bmatrix} I_1 & 0 \\ -F_{21} & I_2 \end{bmatrix} = \begin{bmatrix} I_1 & 0 \\ -C_{21}Q_1 & I_2 \end{bmatrix}$$

la cual se toma como matriz inicial en el proceso iterativo. Para el modelo lineal en tres etapas esta matriz de transformación esta dada por:

$$F_3 = \begin{bmatrix} & & 0 \\ & F_2 & \\ -F_{31} & -F_{32} & I_3 \end{bmatrix} = \begin{bmatrix} I_1 & 0 & 0 \\ -F_{21} & I_2 & 0 \\ -F_{31} & -F_{32} & I_3 \end{bmatrix}$$

En F_3 , los términos F_{31} y F_{32} se determinan a partir de F_2 de la siguiente manera:

$$\begin{aligned} F_{32} &= C_{32}Q_2 \\ F_{31} &= C_{31}Q_1 - C_{32}Q_2C_{21}Q_1 \\ &= C_{31}Q_1 + (-F_{32}) * C_{21}Q_1 \\ &= C_{31}Q_1 + C_{32}Q_2 * (-F_{21}) \end{aligned}$$

De la misma manera, para un modelo lineal en cuatro etapas se tiene

$$F_4 = \begin{bmatrix} & & & \mathbf{0} \\ & F_3 & & \mathbf{0} \\ & & & \mathbf{0} \\ -F_{41} & -F_{42} & -F_{43} & I_4 \end{bmatrix} = \begin{bmatrix} I_1 & 0 & 0 & 0 \\ -F_{21} & I_2 & 0 & 0 \\ -F_{31} & -F_{32} & I_3 & 0 \\ -F_{41} & -F_{42} & -F_{43} & I_4 \end{bmatrix}$$

y los términos F_{41} , F_{42} , F_{43} se obtienen de F_3 de la misma manera como se procedió para encontrar los términos correspondientes a la etapa tres en el modelo de tres etapas. En este sentido se obtiene:

$$\begin{aligned} F_{43} &= C_{43}Q_3 \\ F_{42} &= C_{42}Q_2 - C_{43}Q_3C_{32}Q_2 \\ &= C_{42}Q_2 + (-F_{43}) * C_{32}Q_2 \\ &= C_{42}Q_2 + C_{43}Q_3 * (-F_{32}) \\ F_{41} &= C_{41}Q_1 - C_{42}Q_2C_{21}Q_1 - C_{43}Q_3C_{31}Q_1 + C_{43}Q_3C_{32}Q_2C_{21}Q_1 \\ &= C_{41}Q_1 - (C_{42}Q_2 - C_{43}Q_3C_{32}Q_2)C_{21}Q_1 - C_{43}Q_3C_{31}Q_1 \\ &= C_{41}Q_1 - C_{42}Q_2C_{21}Q_1 - C_{43}Q_3(C_{31}Q_1 - C_{32}Q_2C_{21}Q_1) \\ &= C_{41}Q_1 + (-F_{42})C_{21}Q_1 + (-F_{43})C_{31}Q_1 \\ &= C_{41}Q_1 + C_{42}Q_2 * (-F_{21}) + C_{43}Q_3 * (-F_{31}) \end{aligned}$$

Siguiendo con el mismo procedimiento para el modelo en 5-etapas regular, la matriz que se obtiene esta dada por:

$$F_5 = \begin{bmatrix} & & & & \mathbf{0} \\ & & & & \mathbf{0} \\ & F_4 & & & \mathbf{0} \\ & & & & \mathbf{0} \\ -F_{51} & -F_{52} & -F_{53} & -F_{54} & I_5 \end{bmatrix} = \begin{bmatrix} I_1 & 0 & 0 & 0 & 0 \\ -F_{21} & I_2 & 0 & 0 & 0 \\ -F_{31} & -F_{32} & I_3 & 0 & 0 \\ -F_{41} & -F_{42} & -F_{43} & I_4 & 0 \\ -F_{51} & -F_{52} & -F_{53} & -F_{54} & I_5 \end{bmatrix}$$

donde los términos F_{5j} para $j = 1, 2, 3, 4$ se obtienen a partir de las expresiones:

$$\begin{aligned}
F_{54} &= C_{54}Q_4 \\
F_{53} &= C_{53}Q_3 - C_{54}Q_4C_{43}Q_3 \\
&= C_{53}Q_3 + (-F_{54})C_{43}Q_3 \\
&= C_{53}Q_3 + C_{54}Q_4(-F_{43}) \\
F_{52} &= C_{52}Q_2 - C_{53}Q_3C_{32}Q_2 - C_{54}Q_4C_{42}Q_2 + C_{54}Q_4C_{43}Q_3C_{32}Q_2 \\
&= C_{52}Q_2 - (C_{53}Q_3 - C_{54}Q_4C_{43}Q_3)C_{32}Q_2 - C_{54}Q_4C_{42}Q_2 \\
&= C_{52}Q_2 - C_{53}Q_3C_{32}Q_2 - C_{54}Q_4(C_{42}Q_2 - C_{43}Q_3C_{32}Q_2) \\
&= C_{52}Q_2 + (-F_{53})C_{32}Q_2 + (-F_{54})C_{42}Q_2 \\
&= C_{52}Q_2 + C_{53}Q_3 * (-F_{32}) + C_{54}Q_4 * (-F_{42}) \\
F_{51} &= C_{51}Q_1 - C_{52}Q_2C_{21}Q_1 - C_{53}Q_3C_{31}Q_1 + C_{53}Q_3C_{32}Q_2C_{21}Q_1 - C_{54}Q_4C_{41}Q_1 + \\
&\quad C_{54}Q_4C_{42}Q_2C_{21}Q_1 + C_{54}Q_4C_{43}Q_3C_{31}Q_1 - C_{54}Q_4C_{43}Q_3C_{32}Q_2C_{21}Q_1 \\
&= C_{51}Q_1 - (C_{52}Q_2 - C_{53}Q_3C_{32}Q_2 - C_{54}Q_4C_{42}Q_2 + C_{54}Q_4C_{43}Q_3C_{32}Q_2)C_{21}Q_1 \\
&\quad - (C_{53}Q_3 - C_{54}Q_4C_{43}Q_3)C_{31}Q_1 - C_{54}Q_4C_{41}Q_1 \\
&= C_{51}Q_1 - C_{52}Q_2C_{21}Q_1 - C_{53}Q_3(C_{31}Q_1 - C_{32}Q_2C_{21}Q_1) - \\
&\quad C_{54}Q_4(C_{41}Q_1 - C_{42}Q_2C_{21}Q_1 - C_{43}Q_3C_{31}Q_1 + C_{43}Q_3C_{32}Q_2C_{21}Q_1) \\
&= C_{51}Q_1 + (-F_{52})C_{21}Q_1 + (-F_{53})C_{31}Q_1 + (-F_{54})C_{41}Q_1 \\
&= C_{51}Q_1 + C_{52}Q_2 * (-F_{21}) + C_{53}Q_3 * (-F_{31}) + C_{54}Q_4 * (-F_{41})
\end{aligned}$$

De los resultados anteriores se aprecia que los elementos $F_{ij}, j = 1, 2, \dots, i-1$ correspondientes a la etapa i pueden obtenerse fácilmente de dos maneras diferentes: La primera como se mencionó antes, a partir de los términos F_{ikj} , para $k < j < i-1$ de la misma etapa i previamente calculados y de las matrices C_{ij} de etapas anteriores o bien a partir de los términos F_{ij} obtenidos en las etapas previas y los términos C_{ij} de la etapa i . En otras palabras, los términos F_{ij} pueden obtenerse a partir de cualquiera de las expresiones siguientes:

$$F_{ij} = C_{ij}Q_j - \sum_{k=j+1}^{i-1} F_{ik}C_{kj}Q_j; \quad i = 2, \dots, p; \quad j = 1, \dots, i-1; \quad i > j. \quad (5.6)$$

$$F_{ij} = C_{ij}Q_j - \sum_{k=j+1}^{i-1} C_{ik}Q_k F_{kj}; i = 2, \dots, p; j = 1, \dots, i-1; i > j. \quad (5.7)$$

De las expresiones anteriores, la (5.6) se demuestra más adelante en el Teorema 35, mientras la equivalencia entre ellas se demuestra en el Teorema 36.

Definición 34 Sea el modelo $y = X\beta + \varepsilon$ cuya estructura está definida por:

$$\left\{ \begin{array}{l} y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_p \end{bmatrix}, X = \begin{bmatrix} X_1 & 0 & 0 & \cdots & 0 & 0 \\ C_{2,1} & X_2 & 0 & \cdots & 0 & 0 \\ C_{3,1} & C_{3,2} & X_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \ddots & \vdots \\ C_{p,1} & C_{p,2} & C_{p,3} & \cdots & C_{p,p-1} & X_p \end{bmatrix}, \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \\ \beta_p \end{bmatrix} \\ \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_p \end{bmatrix}, V_\varepsilon = \begin{bmatrix} \sigma_1^2 V_1 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 V_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_p^2 V_p \end{bmatrix} \end{array} \right.$$

donde $Cov(\varepsilon) = V_\varepsilon = \text{Diag}(\sigma_h^2 V_h)$, los vectores $y_1, y_2, \dots, y_p$ son de $n_h \times 1$, para $h = 1, \dots, p$ respectivamente; las matrices $X_1, X_2, \dots, X_p$ y las matrices de enlace $C_{i,j}$ de la etapa i con la etapa j , para $i = 2, \dots, p$, $j = 1, \dots, i-1$, $i > j$, son conocidas de números reales; $\beta_1, \beta_2, \dots, \beta_p$ vectores de parámetros desconocidos de dimensión $p_h \times 1$; y los vectores de error $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p$ aleatorios y no correlacionados entre si con $E(\varepsilon_h) = 0$, $Cov(\varepsilon_h) = \sigma_h^2 V_h$, donde los $\sigma_h^2 > 0$ son parámetros de covarianza desconocidos y las matrices V_h conocidas de dimensión apropiada. Además, se asume que las matrices X_h son de rango completo por columnas y las V_h son definidas positivas, para $h = 1, \dots, p$. A este modelo se le conoce como un modelo lineal en p etapas regular.

Teorema 35 La matriz de transformación F_p del modelo lineal en p etapas regular $y = X\beta + \varepsilon$, especificado en la Definición 34, en el modelo transformado en p etapas

$y^* = X^*\beta + \varepsilon^*$, donde $y^* = F_p y$, $X^* = F_p X$, $\varepsilon^* = F_p \varepsilon$ está dada por:

$$F_p = \begin{bmatrix} I_1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ -F_{21} & I_2 & \mathbf{0} & \cdots & \mathbf{0} \\ -F_{31} & -F_{32} & I_3 & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -F_{p1} & -F_{p2} & -F_{p3} & \cdots & I_p \end{bmatrix} \quad (5.8)$$

donde cada elemento F_{ij} de la matriz anterior se obtiene a partir de la expresión

$$F_{ij} = C_{ij}Q_j - \sum_{k=j+1}^{i-1} F_{ik}C_{kj}Q_j; \quad i = 2, \dots, p; \quad j = 1, \dots, i-1; \quad i > j. \quad (5.9)$$

donde X^* es una matriz diagonal por bloques, con las matrices X_h especificadas en la Definición 34, en la diagonal principal, y las submatrices $\mathbf{0}$ e I_h son matrices nulas e identidad respectivamente de dimensiones apropiadas, para $h = 1, 2, \dots, p$.

Demostración. La demostración procede aplicando el método de inducción en tres pasos:

1. Primero se proba que la expresión (5.9) es válida para un modelo lineal en 3-etapas regular.
2. Luego se asume que es válida para un modelo lineal en $p-1$ -etapas.
3. Finalmente se prueba que es válida también para un modelo lineal en p -etapas

Específicamente lo que se debe probar es que $F^i X = X^*$ para cualquier modelo lineal en i -etapas, $i = 1, 2, \dots, p$.

1. Sea $i = 3$. Entonces las matrices X del modelo y de transformación F_3 son:

$$X = \begin{bmatrix} X_1 & 0 & 0 \\ C_{21} & X_2 & 0 \\ C_{31} & C_{32} & X_3 \end{bmatrix}; \quad F_3 = \begin{bmatrix} I_1 & 0 & 0 \\ -F_{21} & I_2 & 0 \\ -F_{31} & -F_{32} & I_3 \end{bmatrix}$$

luego

$$\begin{bmatrix} I_1 & 0 & 0 \\ -F_{21} & I_2 & 0 \\ -F_{31} & -F_{32} & I_3 \end{bmatrix} \begin{bmatrix} X_1 & 0 & 0 \\ C_{21} & X_2 & 0 \\ C_{31} & C_{32} & X_3 \end{bmatrix} = \begin{bmatrix} X_1 & 0 & 0 \\ -F_{21}X_1 + C_{21} & X_2 & 0 \\ -F_{31}X_1 - F_{32}C_{21} + C_{31} & -F_{32}X_2 + C_{32} & X_3 \end{bmatrix}$$

Del resultado obtenido se sigue:

$$\begin{aligned} -F_{21}X_1 + C_{21} &= [-F_{21} + C_{21}Q_1]X_1 \\ &= [-F_{21} + F_{21}]X_1 \\ &= [0]X_1 = 0 \end{aligned}$$

$$\begin{aligned} -F_{31}X_1 - F_{32}C_{21} + C_{31} &= [-F_{31} - F_{32}C_{21}Q_1 + C_{31}Q_1]X_1 \\ &= [-F_{31} + (C_{31}Q_1 - F_{32}C_{21}Q_1)]X_1 \\ &= [-F_{31} + F_{31}]X_1 \\ &= [0]X_1 = 0 \end{aligned}$$

$$\begin{aligned} -F_{32}X_2 + C_{32} &= [-F_{32} + C_{32}Q_2]X_2 \\ &= [-F_{32} + F_{32}]X_2 \\ &= [0]X_2 = 0 \end{aligned}$$

En consecuencia,

$$\begin{bmatrix} I_1 & 0 & 0 \\ -F_{21} & I_2 & 0 \\ -F_{31} & -F_{32} & I_3 \end{bmatrix} \begin{bmatrix} X_1 & 0 & 0 \\ C_{21} & X_2 & 0 \\ C_{31} & C_{32} & X_3 \end{bmatrix} = \begin{bmatrix} X_1 & 0 & 0 \\ 0 & X_2 & 0 \\ 0 & 0 & X_3 \end{bmatrix}$$

2. Asumiendo que la expresión (5.9) es válida para un modelo en $p - 1$ etapas se

cumple:

$$\begin{aligned}
 & \begin{bmatrix} I_1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ -F_{21} & I_2 & \mathbf{0} & \cdots & \mathbf{0} \\ -F_{31} & -F_{32} & I_3 & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -F_{(p-1)1} & -F_{(p-1)2} & -F_{(p-1)3} & \cdots & I_{(p-1)} \end{bmatrix} \begin{bmatrix} X_1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ C_{21} & X_2 & \mathbf{0} & \cdots & \mathbf{0} \\ C_{31} & C_{32} & X_3 & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ C_{(p-1)1} & C_{(p-1)2} & C_{(p-1)3} & \cdots & X_{(p-1)} \end{bmatrix} \\
 &= \begin{bmatrix} X_1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & X_2 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & X_3 & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & X_{(p-1)} \end{bmatrix}
 \end{aligned}$$

3. Ahora se muestra que la expresión (5.9) es válida también para un modelo lineal en p-etapas. Esto es, que al obtener los términos F_{ij} con la expresión (5.9) se cumple:

$$\begin{aligned}
 & F_p X = X^* = \\
 & \begin{bmatrix} I_1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ -F_{21} & I_2 & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ -F_{31} & -F_{32} & I_3 & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -F_{(p-1)1} & -F_{(p-1)2} & -F_{(p-1)3} & \cdots & I_{(p-1)} & \mathbf{0} \\ -F_{p1} & -F_{p2} & -F_{p3} & \cdots & -F_{p(p-1)} & I_p \end{bmatrix} \begin{bmatrix} X_1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ C_{21} & X_2 & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ C_{31} & C_{32} & X_3 & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ C_{(p-1)1} & C_{(p-1)2} & C_{(p-1)3} & \cdots & X_{(p-1)} & \mathbf{0} \\ C_{p1} & C_{p2} & C_{p3} & \cdots & C_{p(p-1)} & X_p \end{bmatrix} \\
 &= \begin{bmatrix} X_1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & X_2 & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & X_3 & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & X_{(p-1)} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & X_p \end{bmatrix}
 \end{aligned}$$

Para ello basta probar que la última fila de F_p por la matriz X da como resultado

la matriz $[\mathbf{0}, \mathbf{0}, \mathbf{0}, \dots, X_p]$. Específicamente se debe mostrar que:

$$-F_{pj}X_j + \sum_{k=j+1}^{p-1} (-F_{pk}C_{kj}) + C_{pj} = \mathbf{0} \text{ para } j = 1, 2, 3, \dots, p-1$$

para ello basta con observar que para $j = 1, 2, 3, \dots, p-1$ se cumple:

$$\begin{aligned} -F_{pj}X_j + \sum_{k=j+1}^{p-1} (-F_{pk}C_{kj}) + C_{pj} &= -F_{pj}X_j - \sum_{k=j+1}^{p-1} F_{pk}C_{kj} + C_{pj} \\ &= \{-F_{pj} - \sum_{k=j+1}^{p-1} F_{pk}C_{kj}Q_j + C_{pj}Q_j\}X_j \\ &= \{-F_{pj} + [C_{pj}Q_j - \sum_{k=j+1}^{p-1} F_{pk}C_{kj}Q_j]\}X_j \end{aligned}$$

pero como según la ecuación (5.9) se cumple:

$$F_{pj} = C_{pj}Q_j - \sum_{k=j+1}^{p-1} F_{pk}C_{kj}Q_j; j = 1, \dots, p-1;$$

se sigue que

$$-F_{pj}X_j + \sum_{k=j+1}^{p-1} (-F_{pk}C_{kj}) + C_{pj} = -F_{pj} + F_{pj} = \mathbf{0} \text{ para } j = 1, 2, 3, \dots, p-1$$

con lo cual se termina la demostración. ■

5.2.1. Equivalencia de las expresiones para obtener F_p

Antes de proceder a demostrar la equivalencia de las expresiones (5.6) y (5.7) por el método de inducción, es necesario definir las siguientes matrices:

$$\mathfrak{C}_i = \begin{bmatrix} C_{21} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ C_{31} & C_{32} & \mathbf{0} & \cdots & \mathbf{0} \\ C_{41} & C_{42} & C_{43} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ C_{i1} & C_{i2} & C_{i3} & \cdots & C_{i(i-1)} \end{bmatrix}; \mathfrak{C}_{qi} = \begin{bmatrix} C_{21}Q_1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ C_{31}Q_1 & C_{32}Q_2 & \mathbf{0} & \cdots & \mathbf{0} \\ C_{41}Q_1 & C_{42}Q_2 & C_{43}Q_3 & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ C_{i1}Q_1 & C_{i2}Q_2 & C_{i3}Q_3 & \cdots & C_{i(i-1)}Q_{(i-1)} \end{bmatrix}$$

$$\mathfrak{F}_i = \begin{bmatrix} F_{21} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ F_{31} & F_{32} & \mathbf{0} & \cdots & \mathbf{0} \\ F_{41} & F_{42} & F_{43} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ F_{i1} & F_{i2} & F_{i3} & \cdots & F_{i(i-1)} \end{bmatrix}; \mathfrak{Q}_i = \begin{bmatrix} Q_1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & Q_2 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & Q_3 & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & Q_{(i-1)} \end{bmatrix}$$

$$\mathfrak{A}_p = \begin{bmatrix} I_2 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ -F_{32} & I_3 & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ -F_{42} & -F_{43} & I_4 & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -F_{(i-1)2} & -F_{(i-1)3} & -F_{(i-1)4} & \cdots & I_{(i-1)} & \mathbf{0} \\ -F_{i2} & -F_{i3} & -F_{i4} & \cdots & -F_{i(i-1)} & I_i \end{bmatrix}$$

para $i = 2, 3, 4, \dots, p$.

Observe que para un modelo lineal en p-etapas, la matriz $\mathfrak{C}_p$ contiene los términos de enlace presentes en el modelo, $\mathfrak{F}_p$ contiene los términos de la matriz de transformación F_p , $\mathfrak{Q}_p$ contiene las matrices Q_j que satisfacen las condiciones $Q_j X_j = I_{p_j}$, $j = 1, 2, \dots, p-1$, y además se cumple $\mathfrak{C}_{qp} = \mathfrak{C}_p \mathfrak{Q}_p$ y $\mathfrak{F}_i = \mathfrak{A}_i \mathfrak{C}_{qi}$.

En particular para cuatro etapas, $i = 4$ y se obtiene:

$$\mathfrak{C}_4 = \begin{bmatrix} C_{21} & \mathbf{0} & \mathbf{0} \\ C_{31} & C_{32} & \mathbf{0} \\ C_{41} & C_{42} & C_{43} \end{bmatrix}; \mathfrak{C}_{q4} = \begin{bmatrix} C_{21}Q_1 & \mathbf{0} & \mathbf{0} \\ C_{31}Q_1 & C_{32}Q_2 & \mathbf{0} \\ C_{41}Q_1 & C_{42}Q_2 & C_{43}Q_3 \end{bmatrix}$$

$$\mathfrak{F}_4 = \begin{bmatrix} F_{21} & \mathbf{0} & \mathbf{0} \\ F_{31} & F_{32} & \mathbf{0} \\ F_{41} & F_{42} & F_{43} \end{bmatrix}; \mathfrak{Q}_4 = \begin{bmatrix} Q_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & Q_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & Q_3 \end{bmatrix}; \mathfrak{A}_4 = \begin{bmatrix} I_2 & \mathbf{0} & \mathbf{0} \\ -F_{32} & I_3 & \mathbf{0} \\ -F_{42} & -F_{43} & I_4 \end{bmatrix}$$

Obsérvese que

$$\begin{aligned} \mathfrak{A}_4 \mathfrak{C}_{q4} &= \begin{bmatrix} I_2 & \mathbf{0} & \mathbf{0} \\ -F_{32} & I_3 & \mathbf{0} \\ -F_{42} & -F_{43} & I_4 \end{bmatrix} \begin{bmatrix} C_{21}Q_1 & \mathbf{0} & \mathbf{0} \\ C_{31}Q_1 & C_{32}Q_2 & \mathbf{0} \\ C_{41}Q_1 & C_{42}Q_2 & C_{43}Q_3 \end{bmatrix} \\ &= \begin{bmatrix} C_{21}Q_1 & \mathbf{0} & \mathbf{0} \\ C_{31}Q_1 - F_{32}C_{21}Q_1 & C_{32}Q_2 & \mathbf{0} \\ C_{41}Q_1 - F_{42}C_{21}Q_1 - F_{43}C_{31}Q_1 & C_{42}Q_2 - F_{43}C_{32}Q_2 & C_{43}Q_3 \end{bmatrix} \\ &= \begin{bmatrix} F_{21} & \mathbf{0} & \mathbf{0} \\ F_{31} & F_{32} & \mathbf{0} \\ F_{41} & F_{42} & F_{43} \end{bmatrix} \\ &= \mathfrak{F}_4 \end{aligned}$$

$$\begin{aligned} \mathfrak{C}_{q4} F_3 &= \begin{bmatrix} C_{21}Q_1 & \mathbf{0} & \mathbf{0} \\ C_{31}Q_1 & C_{32}Q_2 & \mathbf{0} \\ C_{41}Q_1 & C_{42}Q_2 & C_{43}Q_3 \end{bmatrix} \begin{bmatrix} I_1 & \mathbf{0} & \mathbf{0} \\ -F_{21} & I_2 & \mathbf{0} \\ -F_{31} & -F_{32} & I_3 \end{bmatrix} \\ &= \begin{bmatrix} C_{21}Q_1 & \mathbf{0} & \mathbf{0} \\ C_{31}Q_1 - C_{32}Q_2 F_{21} & C_{32}Q_2 & \mathbf{0} \\ C_{41}Q_1 - C_{42}Q_2 F_{21} - C_{43}Q_3 F_{31} & C_{42}Q_2 - C_{43}Q_3 F_{32} & C_{43}Q_3 \end{bmatrix} \\ &= \mathfrak{T}_4 \end{aligned} \tag{5.10}$$

Teorema 36 La expresión (5.9) empleada para obtener las submatrices F_{ij} de la matriz de transformación (5.8) en un modelo en p -etapas para $i = 2, 3, \dots, p-1, j = 1, 2, \dots, i-1, i > j$ es equivalente a la expresión (5.7) dada por

$$F_{ij} = C_{ij}Q_j - \sum_{k=j+1}^{i-1} C_{ik}Q_k F_{kj}; i = 2, \dots, p; j = 1, \dots, i-1; i > j.$$

Demostración. La demostración se llevará a cabo aplicando el principio de inducción. Para ello primero se mostrara que las expresiones son equivalentes para obtener las submatrices F_{ij} , $i = 2, 3, 4; j = 1, 2, 3; i > j$, en un modelo en 4-etapas. Luego se asumirá que son equivalentes para en el caso de un modelo con p-1 etapas, para finalmente demostrar que también son equivalentes si el modelo posee p etapas.

En el caso de un modelo de cuatro etapas, las ecuaciones (5.6) y (5.7) son equivalentes si los términos F_{ij} obtenidos a través de estas expresiones son iguales para cada $i = 2, 3, \dots, 4; j = 1, 2, 3$, o equivalentemente si se cumple que $\mathfrak{F}_4 = \mathfrak{I}_4$. En este sentido, la demostración para el caso de $i = 4$ etapas consiste en mostrar que la igualdad $\mathfrak{F}_4 = \mathfrak{I}_4$ se cumple. El procedimiento a seguir consiste en mostrar que:

$$\mathfrak{C}_{q4} F_3 X_3 = \mathfrak{C}_4 = \mathfrak{A}_4 \mathfrak{C}_{q4} X_3 \Rightarrow \mathfrak{C}_{q4} F_3 = \mathfrak{A}_4 \mathfrak{C}_{q4}$$

y por ello en el caso de $i = 4$ las ecuaciones (5.6) y (5.7) son equivalentes.

De la ecuación (5.10) se sabe

$$\mathfrak{C}_{q4} F_3 = \begin{bmatrix} C_{21}Q_1 & 0 & 0 \\ C_{31}Q_1 - C_{32}Q_2F_{21} & C_{32}Q_2 & 0 \\ C_{41}Q_1 - C_{42}Q_2F_{21} - C_{43}Q_3F_{31} & C_{42}Q_2 - C_{43}Q_3F_{32} & C_{43}Q_3 \end{bmatrix}$$

luego

$$\mathfrak{C}_{q4} F_3 X_3 = \begin{bmatrix} C_{21}Q_1 & 0 & 0 \\ C_{31}Q_1 - C_{32}Q_2F_{21} & C_{32}Q_2 & 0 \\ C_{41}Q_1 - C_{42}Q_2F_{21} - C_{43}Q_3F_{31} & C_{42}Q_2 - C_{43}Q_3F_{32} & C_{43}Q_3 \end{bmatrix} \begin{bmatrix} X_1 & 0 & 0 \\ C_{21} & X_2 & 0 \\ C_{31} & C_{32} & X_3 \end{bmatrix}$$

Al hacer el producto matricial se encuentra:

- $C_{21}Q_1X_1 = C_{21}$
- $[C_{31}Q_1 - C_{32}Q_2F_{21}]X_1 + C_{32}Q_2C_{21} = [C_{31}Q_1 - C_{32}Q_2F_{21} + C_{32}Q_2C_{21}Q_1]X_1 = [C_{31}Q_1 - C_{32}Q_2F_{21} + C_{32}Q_2F_{21}]X_1 = C_{31}$

$$\blacksquare C_{32}Q_2X_2 = C_{32}$$

$$\begin{aligned} \blacksquare & [C_{41}Q_1 - C_{42}Q_2F_{21} - C_{43}Q_3F_{31}]X_1 + [C_{42}Q_2 - C_{43}Q_3F_{32}]C_{21} + C_{43}Q_3C_{31} \\ &= [C_{41}Q_1 - C_{42}Q_2F_{21} - C_{43}Q_3F_{31} + C_{42}Q_2C_{21}Q_1 - C_{43}Q_3F_{32}C_{21}Q_1 + C_{43}Q_3C_{31}Q_1]X_1 \\ &= [C_{41}Q_1 - C_{42}Q_2F_{21} - C_{43}Q_3F_{31} + C_{42}Q_2F_{21} - C_{43}Q_3F_{32}C_{21}Q_1 + C_{43}Q_3C_{31}Q_1]X_1 \\ &= [C_{41}Q_1 - C_{43}Q_3F_{31} - C_{43}Q_3F_{32}C_{21}Q_1 + C_{43}Q_3C_{31}Q_1]X_1 \\ &= [C_{41}Q_1 - C_{43}Q_3F_{31} + C_{43}Q_3\{C_{31}Q_1 - F_{32}C_{21}Q_1\}]X_1 \\ &= [C_{41}Q_1 - C_{43}Q_3F_{31} + C_{43}Q_3F_{31}]X_1 = C_{41} \end{aligned}$$

$$\begin{aligned} \blacksquare & [C_{42}Q_2 - C_{43}Q_3F_{32}]X_2 + C_{43}Q_3C_{32} \\ &= [C_{42}Q_2 - C_{43}Q_3F_{32} + C_{43}Q_3C_{32}Q_2]X_2 \\ &= [C_{42}Q_2 - C_{43}Q_3F_{32} + C_{43}Q_3F_{32}]X_2 = C_{42} \end{aligned}$$

$$\blacksquare C_{43}Q_3X_3 = C_{43}$$

en consecuencia, $\mathfrak{C}_{q4}F_3X_3 = \mathfrak{C}_4$.

Por otra parte se tiene

$$\begin{aligned} \mathfrak{F}_4X_3 &= \begin{bmatrix} F_{21} & \mathbf{0} & \mathbf{0} \\ F_{31} & F_{32} & \mathbf{0} \\ F_{41} & F_{42} & F_{43} \end{bmatrix} \begin{bmatrix} X_1 & \mathbf{0} & \mathbf{0} \\ C_{21} & X_2 & \mathbf{0} \\ C_{31} & C_{32} & X_3 \end{bmatrix} \\ &= \begin{bmatrix} F_{21}X_1 & \mathbf{0} & \mathbf{0} \\ F_{31}X_1 + F_{32}C_{21} & F_{32}X_2 & \mathbf{0} \\ F_{41}X_1 + F_{42}C_{21} + F_{43}C_{31} & F_{42}X_2 + F_{43}C_{32} & F_{43}X_3 \end{bmatrix} \end{aligned}$$

Debido a que se cumple:

$$\blacksquare F_{21}X_1 = C_{21}Q_1X_1 = C_{21}$$

$$\begin{aligned} \blacksquare F_{31}X_1 + F_{32}C_{21} &= \{F_{31} + F_{32}C_{21}Q_1\}X_1 = \{F_{31} + C_{31}Q_1 - [C_{31}Q_1 - F_{32}C_{21}Q_1]\}X_1 = \\ &= \{F_{31} + C_{31}Q_1 - F_{31}\}X_1 = C_{31} \end{aligned}$$

$$\blacksquare F_{32}X_2 = C_{32}Q_2X_2 = C_{32}$$

- $F_{41}X_1 + F_{42}C_{21} + F_{43}C_{31} = \{F_{41} + F_{42}C_{21}Q_1 + F_{43}C_{31}Q_1\}X_1 = \{F_{41} + C_{41}Q_1 - [C_{41}Q_1 - F_{42}C_{21}Q_1 - F_{43}C_{31}Q_1]\}X_1 = \{F_{41} + C_{41}Q_1 - F_{41}\} = C_{41}$
- $F_{42}X_2 + F_{43}C_{32} = \{F_{42} + F_{43}C_{32}Q_2\}X_2 = \{F_{42} + C_{42}Q_2 - [C_{42}Q_2 - F_{43}C_{32}Q_2]\}X_2 = \{F_{42} + C_{42}Q_2 - F_{42}\}X_2 = C_{42}$
- $F_{43}X_3 = C_{43}Q_3X_3 = C_{43}$

se sigue que

$$\mathfrak{F}_4 X_3 = \begin{bmatrix} C_{21} & \mathbf{0} & \mathbf{0} \\ C_{31} & C_{32} & \mathbf{0} \\ C_{41} & C_{42} & C_{43} \end{bmatrix} = \mathfrak{C}_4$$

y dado que $\mathfrak{F}_4 = \mathfrak{A}_4 \mathfrak{C}_{q4}$, al sustituir en la ecuación anterior se obtiene $\mathfrak{F}_4 X_3 = \mathfrak{A}_4 \mathfrak{C}_{q4} X_3 = C_4$. De esta manera se verifica

$$\mathfrak{C}_{q4} F_3 X_3 = \mathfrak{C}_4 = \mathfrak{A}_4 \mathfrak{C}_{q4} X_3$$

o equivalentemente

$$\mathfrak{C}_{q4} F_3 X_3 = \mathfrak{A}_4 \mathfrak{C}_{q4} X_3$$

Sea $u_3 = X_3 t_3$, un vector cualquiera perteneciente al subespacio $\mathcal{M}(X_3)$. Al post-multiplicar la ecuación anterior por t_3 se obtiene:

$$\mathfrak{C}_{q4} F_3 X_3 t_3 = \mathfrak{A}_4 C_{q4} X_3 t_3$$

$$\mathfrak{C}_{q4} F_3 u_3 = \mathfrak{A}_4 C_{q4} u_3$$

En otras palabras, $\mathfrak{C}_{q4} F_3 u_3 = \mathfrak{A}_4 C_{q4} u_3$ para todo vector $u_3 \in \mathcal{M}(X_3)$, en consecuencia de acuerdo al Lema 2.3.2 de Harville (1997) se sigue que $\mathfrak{C}_{q4} F_3 = \mathfrak{A}_4 C_{q4}$, quedando así demostrado que $\mathfrak{F}_4 = \mathfrak{A}_4$ y por lo tanto que las expresiones (5.6) y (5.7) son equivalentes en el caso de un modelo en 4-etapas.

Asumiendo que lo anterior se cumple para el modelo en $p-1$ etapas, a continuación se mostrará que las expresiones (5.6) y (5.7) para $i = 2, 3, \dots, p$, $j = 1, 2, \dots, i-1$, son también válidas para el modelo en p etapas.

Para ello, similar al caso de $i = 4$ etapas se mostrará que

$$\mathfrak{C}_{qp}F_{p-1}X_{p-1} = \mathfrak{A}_p\mathfrak{C}_{qp}X_{p-1} \Rightarrow \mathfrak{C}_{q4}F_3 = \mathfrak{A}_4\mathfrak{C}_{q4}$$

Obsérvese que:

$$\mathfrak{C}_{qp}F_{p-1} =$$

$$\begin{bmatrix} C_{21}Q_1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ C_{31}Q_1 & C_{32}Q_2 & \mathbf{0} & \cdots & \mathbf{0} \\ C_{41}Q_1 & C_{42}Q_2 & C_{43}Q_3 & & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ C_{p1}Q_1 & C_{p2}Q_2 & C_{p3}Q_3 & \cdots & C_{p(p-1)}Q_{(p-1)} \end{bmatrix} \begin{bmatrix} I_1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ -F_{21} & I_2 & \mathbf{0} & \cdots & \mathbf{0} \\ -F_{31} & -F_{32} & I_3 & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -F_{(p-1)1} & -F_{(p-1)2} & -F_{(p-1)3} & \cdots & I_{(p-1)} \end{bmatrix}$$

$$\mathfrak{C}_{qp}F_{p-1} =$$

$$\begin{bmatrix} C_{21}Q_1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ C_{31}Q_1 - C_{32}Q_2F_{21} & C_{32}Q_2 & \mathbf{0} & \cdots & \mathbf{0} \\ C_{41}Q_1 - C_{42}Q_2F_{21} - C_{43}Q_3F_{31} & C_{42}Q_2 - C_{43}Q_3F_{32} & C_{43}Q_3 & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ C_{p1}Q_1 - \sum_{k=2}^{p-1} C_{pk}Q_kF_{k1} & C_{p2}Q_2 - \sum_{k=3}^{p-1} C_{pk}Q_kF_{k2} & C_{p3}Q_3 - \sum_{k=4}^{p-1} C_{pk}Q_kF_{k3} & \cdots & C_{p(p-1)}Q_{(p-1)} \end{bmatrix}$$

Donde cada elemento de $\mathfrak{C}_{qp}F_{p-1}$ está dado por:

$$\mathfrak{C}_{qp}F_{p-1}(i-1, j) = C_{ij}Q_j - \sum_{k=j+1}^{i-1} C_{ik}Q_kF_{kj}; i = 2, \dots, p; j = 1, \dots, i-1; i > j$$

y $\mathfrak{C}_{qp}F_{p-1}(i-1, j)$ una submatriz de $\mathfrak{C}_{qp}F_{p-1}$ ubicada en la posición $(i-1, j)$. En particular, $\mathfrak{C}_{qp}F_{p-1}(1, 1) = C_{21}Q_1$ y ocurre cuando $i = 2$ y $j = 1$. Así se obtiene:

$$\mathfrak{C}_{qp} F_{p-1} X_{p-1} = \begin{bmatrix} C_{21} Q_1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ C_{31} Q_1 - C_{32} Q_2 F_{21} & C_{32} Q_2 & \mathbf{0} & \cdots & \mathbf{0} \\ C_{41} Q_1 - C_{42} Q_2 F_{21} - C_{43} Q_3 F_{31} & C_{42} Q_2 - C_{43} Q_3 F_{32} & C_{43} Q_3 & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ C_{p1} Q_1 - \sum_{k=2}^{p-1} C_{pk} Q_k F_{k1} & C_{p2} Q_2 - \sum_{k=3}^{p-1} C_{pk} Q_k F_{k2} & C_{p3} Q_3 - \sum_{k=4}^{p-1} C_{pk} Q_k F_{k3} & \cdots & C_{p(p-1)} Q_{(p-1)} \end{bmatrix}$$

$$\begin{bmatrix} X_1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ C_{21} & X_2 & \mathbf{0} & \cdots & \mathbf{0} \\ C_{31} & C_{32} & X_3 & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ C_{(p-1)1} & C_{(p-1)2} & C_{(p-1)3} & \cdots & X_{(p-1)} \end{bmatrix}$$

Como se asumió que los resultados obtenidos en el caso de $i = 4$ son válidos para el caso $i = p - 1$, solo falta mostrar que se mantienen para $i = p$. Para ello sólo es necesario multiplicar la última fila de $\mathfrak{C}_{qp} F_{p-1}$ por cada una de las $p - 1$ columnas de X_{p-1} desde la 1 a la $p - 1$, y obtener $C_{p1}, C_{p2}, \dots, C_{p(p-1)}$ respectivamente.

- Con la primera columna de X_{p-1} .

$$\begin{aligned} & \left(C_{p1} Q_1 - \sum_{k=2}^{p-1} C_{pk} Q_k F_{k1} \right) X_1 + \left(C_{p2} Q_2 - \sum_{k=3}^{p-1} C_{pk} Q_k F_{k2} \right) C_{21} + \left(C_{p3} Q_3 - \sum_{k=4}^{p-1} C_{pk} Q_k F_{k3} \right) C_{31} \\ & + \cdots + (C_{p(p-2)} Q_{(p-2)} - C_{p(p-1)} Q_{(p-1)} F_{(p-1)(p-2)}) C_{(p-2)1} + C_{p(p-1)} Q_{(p-1)} C_{(p-1)1} \\ & \left[\left(C_{p1} Q_1 - \sum_{k=2}^{p-1} C_{pk} Q_k F_{k1} \right) + \left(C_{p2} Q_2 - \sum_{k=3}^{p-1} C_{pk} Q_k F_{k2} \right) C_{21} Q_1 + \left(C_{p3} Q_3 - \sum_{k=4}^{p-1} C_{pk} Q_k F_{k3} \right) C_{31} Q_1 \right. \\ & + \cdots + (C_{p(p-2)} Q_{(p-2)} - C_{p(p-1)} Q_{(p-1)} F_{(p-1)(p-2)}) C_{(p-2)1} Q_1 + C_{p(p-1)} Q_{(p-1)} C_{(p-1)1} Q_1 \left. \right] X_1 \\ & [C_{p1} Q_1 - C_{p2} Q_2 F_{21} - C_{p3} Q_3 F_{31} - C_{p4} Q_4 F_{41} - \cdots - C_{p(p-1)} Q_{(p-1)} F_{(p-1)1} + \\ & C_{p2} Q_2 C_{21} Q_1 - C_{p3} Q_3 F_{32} C_{21} Q_1 - C_{p4} Q_4 F_{42} C_{21} Q_1 - \cdots - C_{p(p-1)} Q_{(p-1)} F_{(p-1)2} C_{21} Q_1 + \\ & C_{p3} Q_3 C_{31} Q_1 - C_{p4} Q_4 F_{43} C_{31} Q_1 - \cdots - C_{p(p-1)} Q_{(p-1)} F_{(p-1)3} C_{31} Q_1 \\ & + \cdots + C_{p(p-2)} Q_{(p-2)} C_{(p-2)1} Q_1 - C_{p(p-1)} Q_{(p-1)} F_{(p-1)(p-2)} C_{(p-2)1} Q_1 + \\ & C_{p(p-1)} Q_{(p-1)} C_{(p-1)1} Q_1] X_1 \end{aligned}$$

Ordenando los términos de manera apropiada se obtiene:

$$\begin{aligned}
& \begin{bmatrix} C_{p1}Q_1 & -C_{p2}Q_2F_{21} & -C_{p3}Q_3F_{31} & \cdots & -C_{p(p-1)}Q_{(p-1)}F_{(p-1)1} \\
& +C_{p2}Q_2C_{21}Q_1 & -C_{p3}Q_3F_{32}C_{21}Q_1 & \cdots & -C_{p(p-1)}Q_{(p-1)}F_{(p-1)2}C_{21}Q_1 \\
& & +C_{p3}Q_3C_{31}Q_1 & \cdots & -C_{p(p-1)}Q_{(p-1)}F_{(p-1)3}C_{31}Q_1 \\
& & & \cdots & -C_{p(p-1)}Q_{(p-1)}F_{(p-1)4}C_{41}Q_1 \\
& & & \ddots & \vdots \\
& & & & -C_{p(p-1)}Q_{(p-1)}F_{(p-1)(p-2)}C_{(p-2)1}Q_1 \\
& & & & +C_{p(p-1)}Q_{(p-1)}C_{(p-1)1}Q_1 \end{bmatrix} X_1
\end{aligned}$$

En consecuencia, solo falta sumar los términos por columnas. Esto es:

$$\checkmark -C_{p2}Q_2F_{21} + C_{p2}Q_2C_{21}Q_1 = -C_{p2}Q_2F_{21} + C_{p2}Q_2F_{21} = 0$$

$$\begin{aligned}
\checkmark -C_{p3}Q_3F_{31} - C_{p3}Q_3F_{32}C_{21}Q_1 + C_{p3}Q_3C_{31}Q_1 \\
= C_{p3}Q_3[-F_{31} + C_{31}Q_1 - F_{32}C_{21}Q_1] \\
= C_{p3}Q_3[-F_{31} + F_{31}] = 0
\end{aligned}$$

$$\begin{aligned}
\checkmark -C_{p(p-1)}Q_{(p-1)}F_{(p-1)1} - C_{p(p-1)}Q_{(p-1)}F_{(p-1)2}C_{21}Q_1 - C_{p(p-1)}Q_{(p-1)}F_{(p-1)3}C_{31}Q_1 \\
-C_{p(p-1)}Q_{(p-1)}F_{(p-1)4}C_{41}Q_1 - \cdots - C_{p(p-1)}Q_{(p-1)}F_{(p-1)(p-2)}C_{(p-2)1}Q_1 + C_{p(p-1)}Q_{(p-1)}C_{(p-1)1}Q_1 \\
= C_{p(p-1)}Q_{(p-1)}[-F_{(p-1)1} + C_{(p-1)1}Q_1 - F_{(p-1)2}C_{21}Q_1 - F_{(p-1)3}C_{31}Q_1 - F_{(p-1)4}C_{41}Q_1 - \\
\cdots - F_{(p-1)(p-2)}C_{(p-2)1}Q_1] \\
= C_{p(p-1)}Q_{(p-1)}[-F_{(p-1)1} + F_{(p-1)1}] = 0
\end{aligned}$$

Por lo tanto resulta:

$$\begin{aligned}
& \left[\left(C_{p1}Q_1 - \sum_{k=2}^{p-1} C_{pk}Q_kF_{k1} \right) + \left(C_{p2}Q_2 - \sum_{k=3}^{p-1} C_{pk}Q_kF_{k2} \right) C_{21}Q_1 + \left(C_{p3}Q_3 - \sum_{k=4}^{p-1} C_{pk}Q_kF_{k3} \right) C_{31}Q_1 \right. \\
& + \cdots + \left(C_{p(p-2)}Q_{(p-2)} - C_{p(p-1)}Q_{(p-1)}F_{(p-1)(p-2)} \right) C_{(p-1)1}Q_1 + C_{p(p-1)}Q_{(p-1)}C_{(p-1)1}Q_1 \Big] X_1 \\
& = [C_{p1}Q_1]X_1 = C_{p1}
\end{aligned}$$

- Con la penúltima columna de X_{p-1} .

$$\begin{aligned}
& \{ C_{p(p-2)}Q_{(p-2)} - C_{p(p-1)}Q_{(p-1)}F_{(p-1)(p-2)} \} X_{(p-2)} + C_{p(p-1)}Q_{(p-1)}C_{(p-1)(p-2)} \\
& = \{ C_{p(p-2)}Q_{(p-2)} - C_{p(p-1)}Q_{(p-1)}F_{(p-1)(p-2)} + C_{p(p-1)}Q_{(p-1)}C_{(p-1)(p-2)}Q_{(p-2)} \} X_{(p-2)} \\
& = \{ C_{p(p-2)}Q_{(p-2)} - C_{p(p-1)}Q_{(p-1)}F_{(p-1)(p-2)} + C_{p(p-1)}Q_{(p-1)}F_{(p-1)(p-2)} \} X_{(p-2)} \\
& = C_{p(p-2)}Q_{(p-2)}X_{(p-2)} = C_{p(p-2)}
\end{aligned}$$

- Con la última columna de X_{p-1} :

$$C_{p(p-1)}Q_{(p-1)}X_{(p-1)} = C_{p(p-1)}.$$

Dado que cada término obtenido del producto $\mathfrak{C}_{pp}F_{p-1}X_{p-1}$ da como resultado una matriz

C_{ij} para $i = 2, 3, \dots, p, j = 1, 2, \dots, i - 1$ se concluye que:

$$\mathfrak{C}_{qp} F_{p-1} X_{p-1} = \mathfrak{C}_p.$$

Ahora se mostrará que $\mathfrak{A}_p \mathfrak{C}_{qp} X_{p-1} = \mathfrak{C}_p$. Para ello se obtiene:

$$\mathfrak{A}_p C_{qp} = \begin{bmatrix} I_2 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ -F_{32} & I_3 & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ -F_{42} & -F_{43} & I_4 & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -F_{p2} & -F_{p3} & -F_{p4} & \cdots & -F_{p(p-1)} & I_p \end{bmatrix} \begin{bmatrix} C_{21}Q_1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ C_{31}Q_1 & C_{32}Q_2 & \mathbf{0} & \cdots & \mathbf{0} \\ C_{41}Q_1 & C_{42}Q_2 & C_{43}Q_3 & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ C_{p1}Q_1 & C_{p2}Q_2 & C_{p3}Q_3 & \cdots & C_{p(p-1)}Q_{(p-1)} \end{bmatrix}$$

$$\mathfrak{A}_p C_{qp} = \begin{bmatrix} C_{21}Q_1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ C_{31}Q_1 - F_{32}C_{21}Q_1 & C_{32}Q_2 & \mathbf{0} & \cdots & \mathbf{0} \\ C_{41}Q_1 - F_{42}C_{21}Q_1 - F_{43}C_{31}Q_1 & C_{42}Q_2 - F_{43}C_{32}Q_2 & C_{43}Q_3 & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ C_{p1}Q_1 - \sum_{k=2}^{p-1} F_{pk}C_{k1}Q_1 & C_{p2}Q_2 - \sum_{k=3}^{p-1} F_{pk}C_{k2}Q_2 & C_{p3}Q_3 - \sum_{k=4}^{p-1} F_{pk}C_{k3}Q_3 & \cdots & C_{p(p-1)}Q_{(p-1)} \end{bmatrix}$$

y por lo tanto

$$\mathfrak{A}_p C_{qp} = \mathfrak{F}_p = \begin{bmatrix} F_{21} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ F_{31} & F_{32} & \mathbf{0} & \cdots & \mathbf{0} \\ F_{41} & F_{42} & F_{43} & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ F_{p1} & F_{p2} & F_{p3} & \cdots & F_{p(p-1)} \end{bmatrix}$$

De esta manera,

$$\mathfrak{A}_p C_{qp} X_{p-1} = \begin{bmatrix} F_{21} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ F_{31} & F_{32} & \mathbf{0} & \cdots & \mathbf{0} \\ F_{41} & F_{42} & F_{43} & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ F_{p1} & F_{p2} & F_{p3} & \cdots & F_{p(p-1)} \end{bmatrix} \begin{bmatrix} X_1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ C_{21} & X_2 & \mathbf{0} & \cdots & \mathbf{0} \\ C_{31} & C_{32} & X_3 & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ C_{(p-1)1} & C_{(p-1)2} & C_{(p-1)3} & \cdots & X_{(p-1)} \end{bmatrix}$$

Al igual que antes, se asume que los resultados del caso de $i = 4$ son validos para $i = p - 1$, y sólo es necesario mostrar que la última fila de $\mathfrak{F}_p = \mathfrak{A}_p C_{qp}$ multiplicada por cada una de las

$p - 1$ columnas de X_{p-1} produce las submatrices C_{ij} para $i = 2, 3, \dots, p$ y $j = 1, 2, \dots, i - 1$.

En sentido, los resultados son los siguientes:

- Con la primera columna de X_{p-1} .

$$\begin{aligned} F_{p1}X_1 + \sum_{k=2}^{p-1} F_{pk}C_{k1} &= F_{p1}X_1 + [C_{p1}Q_1 - \{C_{p1}Q_1 - \sum_{k=2}^{p-1} F_{pk}C_{k1}Q_1\}]X_1 \\ &= F_{p1}X_1 + [C_{p1}Q_1 - F_{p1}]X_1 \\ &= F_{p1}X_1 + C_{p1} - F_{p1}X_1 = C_{p1} \\ &\vdots \end{aligned}$$

- Con la penúltima columna de X_{p-1} .

$$\begin{aligned} F_{p(p-2)}X_{p-2} + F_{p(p-1)}C_{(p-1)(p-2)} \\ &= F_{p(p-2)}X_{p-2} + [C_{p(p-2)}Q_{p-2} - \{C_{p(p-2)}Q_{p-2} - F_{p(p-1)}C_{(p-1)(p-2)}Q_{p-2}\}]X_{p-2} \\ &= F_{p(p-2)}X_{p-2} + [C_{p(p-2)}Q_{p-2} - F_{p(p-2)}]X_{p-2} = F_{p(p-2)}X_{p-2} + C_{p(p-2)} - F_{p(p-2)}X_{p-2} \\ &= C_{p(p-2)} \end{aligned}$$

- Con la última columna de X_{p-1} :

$$F_{p(p-1)}X_{p-1} = C_{p(p-1)}Q_{(p-1)}X_{(p-1)} = C_{p(p-1)}.$$

comprobándose de esta manera que $\mathfrak{A}_p C_{qp} X_{p-1} = C_p$.

Sea $u_{t-1} = X_{p-1}t_{p-1}$, un vector cualquiera perteneciente al subespacio $\mathcal{M}(X_{p-1})$. Debido a que $\mathfrak{E}_{qp}F_{p-1}X_{p-1} = \mathfrak{E}_p$ y $\mathfrak{A}_p C_{qp} X_{p-1} = C_p$ se sigue que:

$$\begin{aligned} \mathfrak{E}_{qp}F_{p-1}X_{p-1} &= \mathfrak{A}_p C_{qp} X_{p-1} \\ \mathfrak{E}_{qp}F_{p-1}X_{p-1}t_{p-1} &= \mathfrak{A}_p C_{qp} X_{p-1}t_{p-1} \\ \mathfrak{E}_{qp}F_{p-1}u_{p-1} &= \mathfrak{A}_p C_{qp} u_{p-1} \end{aligned}$$

En otras palabras, $\mathfrak{E}_{qp}F_{p-1}u_{p-1} = \mathfrak{A}_p C_{qp} u_{p-1}$ para todo vector u_{p-1} en $\mathcal{M}(X_{p-1})$, en consecuencia de acuerdo al Lema 2.3.2 de Harville (1997) se sigue que $\mathfrak{E}_{qp}F_{p-1} = \mathfrak{A}_p C_{qp}$, quedando demostrado de esta manera que las expresiones (5.6) y (5.7) son equivalentes para obtener las submatrices F_{ij} de la matriz de transformación en un modelo de p-etapas. ■

5.3. Obtención del estimador UBLUE para β y $X\beta$.

En esta sección se dan expresiones para los estimadores UBLUE del modelo en p etapas $y = X\beta + \varepsilon$.

Sean $\mathcal{M}(C_{ij})$ y $\mathcal{M}(X_i)$ los espacios vectoriales generados por las columnas de las matrices C_{ij} y X_i respectivamente, $i = 2, \dots, p$; $j = 1, \dots, p-1$, donde C_{ij} y X_i , son matrices relacionadas a la etapa i del modelo lineal en p etapas. Supongáse que en cada etapa i del modelo, para $i = 2, \dots, p$, se cumplen las siguientes condiciones:

$$\mathcal{M}(C_{ij}) \subset \mathcal{M}(X_i) \quad \text{para } j = 1, \dots, p-1 \quad (5.11)$$

Dado que cada termino $F_{ij} = C_{ij}Q_j - \sum_{k=j+1}^{i-1} C_{ik}Q_k F_{kj}$ es una suma de productos de matrices donde el primer miembro en cada término es una matriz C_{ij} perteneciente a la etapa i , se sigue a partir de (5.11) que

$$\mathcal{M}(F_{ij}) \subset \mathcal{M}(X_i) \quad \forall i = 2, \dots, p; j = 1, \dots, p-1; i > j \quad (5.12)$$

Sea el modelo transformado en p etapas

$$y^* = X^*\beta + \varepsilon^* \quad (5.13)$$

obtenido al pre-multiplicar la matriz F_p dada en (5.8) por el modelo lineal en p etapas regular $y = X\beta + \varepsilon$ dado en la Definición 34; esto es, $y^* = F_p y$, $X^* = F_p X$, $\varepsilon^* = F_p \varepsilon$. En consecuencia, debido a que $Cov(\varepsilon) = V_\varepsilon = Diag(\sigma_h^2 V_h)$, se sigue que $Cov(\varepsilon^*) = V_{\varepsilon^*} = Cov(F_p \varepsilon) = F_p V_\varepsilon F_p' = F_p Diag(\sigma_h^2 V_h) F_p'$ para $h = 1, \dots, p$.

Del Teorema 35 es claro que $X^* = FX$ es la matriz diagonal en bloques $Diag(X_h)$, $h = 1, \dots, p$. En consecuencia

$$X^* = \begin{bmatrix} X_1 & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & X_2 & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & X_{p-1} & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & X_p \end{bmatrix}$$

$$\begin{aligned}
P_{X^*} &= \begin{bmatrix} P_{X_1} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & P_{X_2} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & P_{X_{p-1}} & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & P_{X_p} \end{bmatrix} \\
Z = I - P_{X^*} &= \begin{bmatrix} Z_1 & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & Z_2 & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & Z_{p-1} & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & Z_p \end{bmatrix} \\
&= \text{Diag}(Z_h)
\end{aligned}$$

donde $P_{X_h} = X_h(X_h'X_h)^-X_h'$ y $Z_h = I_h - P_{X_h}$ para $h = 1, \dots, p$.

Es claro que $F_p'Z = ZF_p = Z$ ya que

$$F_p'Z = \begin{bmatrix} Z_1 & -F_{21}'Z_2 & \cdots & -F_{p-11}'Z_{p-1} & -F_{p1}'Z_p \\ \mathbf{0} & Z_2 & \cdots & -F_{p-12}'Z_{p-1} & -F_{p2}'Z_p \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & Z_{p-1} & -F_{pp-1}'Z_p \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & Z_p \end{bmatrix}$$

y en virtud de (5.12) se sigue que $Z_i F_{ij} = 0$ o equivalentemente que $F_{ij}'Z_i = 0$ para todo $i = 2, \dots, p$ y todo $j = 1, \dots, p-1$; $i > j$.

Teorema 37 *Considérese el modelo lineal regular en p etapas dado en la Definición 34. Si en cada etapa i del modelo las condiciones $\mathcal{M}(C_{ij}) \subset \mathcal{M}(X_i)$ se satisfacen para todo $i = 2, \dots, p$ y todo $j = 1, \dots, p-1$, $i > j$, entonces*

1. El BLUE de $X^*\beta$ obtenido a partir del modelo (5.13) esta dado por

$$\begin{aligned} \widehat{X^*\beta} &= [I - V_{\varepsilon^*} (I - P_{X^*}) \{(I - P_{X^*}) V_{\varepsilon^*} (I - P_{X^*})\}^- (I - P_{X^*})] y^* \\ &= \begin{bmatrix} I_1 - P'_{Z_1 V_1} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ F_{21} P'_{Z_1 V_1} & I_2 - P'_{Z_2 V_2} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ F_{p-11} P'_{Z_1 V_1} & F_{p-12} P'_{Z_2 V_2} & \cdots & I_{p-1} - P'_{Z_{p-1} V_{p-1}} & \mathbf{0} \\ F_{p1} P'_{Z_1 V_1} & F_{p2} P'_{Z_2 V_2} & \cdots & F_{pp-1} P'_{Z_{p-1} V_{p-1}} & I_p - P'_{Z_p V_p} \end{bmatrix} y^* \end{aligned}$$

2. Una representación del BLUE de $X^*\beta$ es

$$\widehat{X^*\beta} = P_{X^*|V_{\varepsilon^*}Z} y^*$$

donde

$$P_{X^*|V_{\varepsilon^*}Z} = \begin{bmatrix} I_1 - P'_{Z_1 V_1} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ F_{21} P'_{Z_1 V_1} & I_2 - P'_{Z_2 V_2} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ F_{p-11} P'_{Z_1 V_1} & F_{p-12} P'_{Z_2 V_2} & \cdots & I_{p-1} - P'_{Z_{p-1} V_{p-1}} & \mathbf{0} \\ F_{p1} P'_{Z_1 V_1} & F_{p2} P'_{Z_2 V_2} & \cdots & F_{pp-1} P'_{Z_{p-1} V_{p-1}} & I_p - P'_{Z_p V_p} \end{bmatrix}$$

es un proyector oblicuo sobre $\mathcal{M}(X^*)$ a lo largo de $\mathcal{M}(V_{\varepsilon^*}Z)$.

3. El BLUE $\widehat{X^*\beta} = P_{X^*|V_{\varepsilon^*}Z} y^*$ es también el UBLUE de $X^*\beta$.

4. Los UBLUE de β y $X\beta$ para el modelo dado en la Definición (34) son

$$\widehat{\beta} = X^{*-} P_{X^*|V_{\varepsilon^*}Z} F y$$

$$\widehat{X\beta} = X X^{*-} P_{X^*|V_{\varepsilon^*}Z} F y$$

5. El proyector oblicuo sobre $\mathcal{M}(X)$ paralelo a (o a lo largo de) $\mathcal{M}(V_{\varepsilon}Z)$ es

$$X X^{*-} P_{X^*|V_{\varepsilon^*}Z} F.$$

Demostración.

1. El BLUE de $X^*\beta$ para el modelo (5.13) se obtiene a partir de la expresión

$$\widehat{X^*\beta} = [I - V_{\varepsilon^*}(I - P_{X^*})]\{(I - P_{X^*})V_{\varepsilon^*}(I - P_{X^*})\}^-(I - P_{X^*})y^*$$

Al hacer algunos cálculos sencillos se aprecia:

$$\begin{aligned} V_{\varepsilon^*}(I - P_{X^*}) &= V_{\varepsilon^*}Z = F_p V_{\varepsilon} F_p' Z \\ &= F_p V_{\varepsilon} Z = F_p \text{Diag}(\sigma_h^2 V_h Z_h) \\ &= \begin{bmatrix} \sigma_1^2 V_1 Z_1 & \mathbf{0} & \cdots & \mathbf{0} \\ -\sigma_1^2 F_{21} V_1 Z_1 & \sigma_2^2 V_2 Z_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ -\sigma_1^2 F_{p1} V_1 Z_1 & -\sigma_2^2 F_{p2} V_2 Z_2 & \cdots & \sigma_p^2 V_p Z_p \end{bmatrix} \end{aligned} \quad (5.14)$$

ya que $F_p' Z = Z$ y $V_{\varepsilon} Z = \text{Diag}(\sigma_h^2 V_h Z_h)$, para $h = 1, \dots, p$. Luego

$$\begin{aligned} (I - P_{X^*})V_{\varepsilon^*}(I - P_{X^*}) &= ZV_{\varepsilon^*}Z = ZF_p \text{Diag}(\sigma_h^2 V_h Z_h) \\ &= Z \text{Diag}(\sigma_h^2 V_h Z_h) = \text{Diag}(Z_h) \text{Diag}(\sigma_h^2 V_h Z_h) \\ &= \text{Diag}(\sigma_h^2 Z_h V_h Z_h) \end{aligned}$$

y por lo tanto

$$\{(I - P_{X^*})V_{\varepsilon^*}(I - P_{X^*})\}^- = \text{Diag}\left(\frac{1}{\sigma_i^2} \{Z_h V_h Z_h\}^-\right)$$

Específicamente,

$$\begin{aligned} (I - P_{X^*})V_{\varepsilon^*}(I - P_{X^*}) &= \text{Diag}(\sigma_h^2 Z_h V_h Z_h) \\ &= \begin{bmatrix} \sigma_1^2 Z_1 V_1 Z_1 & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ 0 & \sigma_2^2 Z_2 V_2 Z_2 & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \sigma_p^2 Z_p V_p Z_p \end{bmatrix} \end{aligned}$$

$$\begin{aligned} \{(I - P_{X^*}) V_{\varepsilon^*} (I - P_{X^*})\}^- &= \text{Diag} \left(\frac{1}{\sigma_h^2} \{Z_h V_h Z_h\}^- \right) \\ &= \begin{bmatrix} \frac{1}{\sigma_1^2} \{Z_1 V_1 Z_1\}^- & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \frac{1}{\sigma_2^2} \{Z_2 V_2 Z_2\}^- & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \frac{1}{\sigma_p^2} \{Z_p V_p Z_p\}^- \end{bmatrix} \end{aligned}$$

De esta manera

$$\begin{aligned} \widehat{X^* \beta} &= [I - V_{\varepsilon^*} (I - P_{X^*}) \{(I - P_{X^*}) V_{\varepsilon^*} (I - P_{X^*})\}^- (I - P_{X^*})] y^* \\ &= [I - F_p \text{Diag} (\sigma_h^2 V_h Z_h) \text{Diag} \left(\frac{1}{\sigma_h^2} \{Z_h V_h Z_h\}^- \right) \text{Diag}(Z_h)] y^* \\ &= [I - F_p \text{Diag} (V_h Z_h \{Z_h V_h Z_h\}^-) \text{Diag}(Z_h)] y^* \\ &= [I - F_p \text{Diag} (V_h Z_h \{Z_h V_h Z_h\}^- Z_h)] y^* \end{aligned}$$

para $h = 1, \dots, p$. Luego, denominando $P'_{Z_h V_h} = Z_h \{Z_h V_h Z_h\}^- Z_h V_h$, se obtiene

$$\begin{aligned} \widehat{X^* \beta} &= [I - V_{\varepsilon^*} (I - P_{X^*}) \{(I - P_{X^*}) V_{\varepsilon^*} (I - P_{X^*})\}^- (I - P_{X^*})] y^* \\ &= [I - F_p \text{Diag} (P'_{Z_h V_h})] y^* \\ &= \begin{bmatrix} I_1 - P'_{Z_1 V_1} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ F_{21} P'_{Z_1 V_1} & I_2 - P'_{Z_2 V_2} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ F_{p-11} P'_{Z_1 V_1} & F_{p-12} P'_{Z_2 V_2} & \cdots & I_{p-1} - P'_{Z_{p-1} V_{p-1}} & \mathbf{0} \\ F_{p1} P'_{Z_1 V_1} & F_{p2} P'_{Z_2 V_2} & \cdots & F_{pp-1} P'_{Z_{p-1} V_{p-1}} & I_p - P'_{Z_p V_p} \end{bmatrix} y^* \end{aligned}$$

2. En esta parte es importante observar que

$$\begin{aligned} [I - F_p \text{Diag} (P'_{Z_h V_h})] X^* &= X^* - F_p \text{Diag} (P'_{Z_h V_h}) X^* \\ &= X^* - F_p \text{Diag} (P'_{Z_h Z_h}) \text{Diag}(X_h) \\ &= X^* - F_p \text{Diag} (P'_{Z_h V_h} X_h) \\ &= X^* - F_p \text{Diag} (V_h Z_h \{Z_h V_h Z_h\}^- Z_h X_h) \\ &= X^* \end{aligned} \tag{5.15}$$

ya que $Z_h X_h = 0$ al ser $\mathcal{M}(X_h) \perp \mathcal{M}(Z_h)$ para todo $h = 1, \dots, p$. Por otra parte de (5.14) se sabe que $V_{\varepsilon^*} Z = F_p \text{Diag}(\sigma_h^2 V_h Z_h)$, de aquí resulta

$$\begin{aligned}
 & \text{Diag}(P'_{Z_h V_h}) V_{\varepsilon^*} Z \\
 = & \text{Diag}(P'_{Z_h V_h}) F_p \text{Diag}(\sigma_h^2 V_h Z_h) \\
 = & \begin{bmatrix} P'_{Z_1 V_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & P'_{Z_2 V_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & P'_{Z_p V_p} \end{bmatrix} \begin{bmatrix} \sigma_1^2 V_1 Z_1 & \mathbf{0} & \cdots & \mathbf{0} \\ -\sigma_1^2 F_{21} V_1 Z_1 & \sigma_2^2 V_2 Z_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ -\sigma_1^2 F_{p1} V_1 Z_1 & -\sigma_2^2 F_{p2} V_2 Z_2 & \cdots & \sigma_p^2 V_p Z_p \end{bmatrix} \\
 = & \begin{bmatrix} \sigma_1^2 P'_{Z_1 V_1} V_1 Z_1 & \mathbf{0} & \cdots & \mathbf{0} \\ -\sigma_1^2 P'_{Z_2 V_2} F_{21} V_1 Z_1 & \sigma_2^2 P'_{Z_2 V_2} V_2 Z_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ -\sigma_1^2 P'_{Z_p V_p} F_{p1} V_1 Z_1 & -\sigma_2^2 P'_{Z_p V_p} F_{p2} V_2 Z_2 & \cdots & \sigma_p^2 P'_{Z_p V_p} V_p Z_p \end{bmatrix}
 \end{aligned}$$

Luego de (5.12) se sigue que para $i = 2, 3, \dots, p$ y $j = 1, 2, \dots, p-1$, $i > j$ $\mathcal{M}(Z_i) \perp \mathcal{M}(F_{ij})$, por lo que $P'_{Z_i V_i} F_{ij} = V_i Z_i \{Z_i V_i Z_i\}^{-1} Z_i F_{ij} = 0$, y de (4.2) se obtiene que para $h = 1, 2, \dots, p$, $P'_{Z_h V_h} V_h Z_h = V_h Z_h$; de aquí que $\text{Diag}(P'_{Z_h V_h}) V_{\varepsilon^*} Z = \text{Diag}(\sigma_h^2 V_h Z_h)$, y entonces.

$$\begin{aligned}
 [I - F_p \text{Diag}(P'_{Z_h V_h})] V_{\varepsilon^*} Z &= V_{\varepsilon^*} Z - F_p \text{Diag}(P'_{Z_h V_h}) V_{\varepsilon^*} Z \\
 &= V_{\varepsilon^*} Z - F_p \text{Diag}(\sigma_h^2 V_h Z_h) \\
 &= V_{\varepsilon^*} Z - V_{\varepsilon^*} Z \\
 &= 0
 \end{aligned} \tag{5.16}$$

por lo que de (5.15) y (5.16) se concluye que $I - F_p \text{Diag}(P'_{Z_h V_h})$ es un proyector oblicuo sobre $\mathcal{M}(X^*)$ paralelo a (o a lo largo de) $\mathcal{M}(V_{\varepsilon^*} Z)$, y de esta forma, el estimador de $X^* \beta$ puede expresarse como:

$$\widehat{X^* \beta} = P_{X^* | V_{\varepsilon^*} Z} y^*$$

donde

$$P_{X^*|V_{\varepsilon^*}Z} = \begin{bmatrix} I_1 - P'_{Z_1V_1} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ F'_{21}P'_{Z_1V_1} & I_2 - P'_{Z_2V_2} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ F'_{p-11}P'_{Z_1V_1} & F'_{p-12}P'_{Z_2V_2} & \cdots & I_{p-1} - P'_{Z_{p-1}V_{p-1}} & \mathbf{0} \\ F'_{p1}P'_{Z_1V_1} & F'_{p2}P'_{Z_2V_2} & \cdots & F'_{pp-1}P'_{Z_{p-1}V_{p-1}} & I_p - P'_{Z_pV_p} \end{bmatrix}$$

3. Dado que $\widehat{X^*\beta} = P_{X^*|V_{\varepsilon^*}Z}y^*$ es un BLUE que no depende en forma explícita de los parámetros de varianza σ_i^2 , $i = 1, \dots, p$, se sigue de Wulff y Birkes (2005) que $\widehat{X^*\beta}$ también es el UBLUE para $X^*\beta$.

4. Como $y^* = Fy$ podemos escribir a $\widehat{X^*\beta}$ en función de Fy ; esto es,

$$\widehat{X^*\beta} = P_{X^*|V_{\varepsilon^*}Z}Fy.$$

Por otra parte, dado que las matrices X_i son de rango completo por columnas, denotando con X^{*-} a la matriz

$$X^{*-} = \begin{bmatrix} (X'_1X_1)^{-1}X'_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & (X'_2X_2)^{-1}X'_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & (X'_pX_p)^{-1}X'_p \end{bmatrix}$$

se sigue que los estimadores de β y de $X\beta$ para el modelo dado en la Definición 34 están dados por

$$\begin{aligned} \beta &= X^{*-}\widehat{X^*\beta} = X^{*-}P_{X^*|V_{\varepsilon^*}Z}Fy \\ \widehat{X\beta} &= XX^{*-}\widehat{X^*\beta} = XX^{*-}P_{X^*|V_{\varepsilon^*}Z}Fy \end{aligned}$$

los cuales son UBLUE por ser funciones lineales del UBLUE de $\widehat{X^*\beta}$ del modelo transformado.

5. Dado que $Z = F'_pZ$ y F_p^{-1} existe por ser las submatrices en la diagonal principal de F_p matrices identidad se cumple:

$$a) XX^{*-}P_{X^*|V_{\varepsilon^*}Z}F'X = XX^{*-}P_{X^*|V_{\varepsilon^*}Z}X^* = XX^{*-}X^* = F_p^{-1}X^*X^{*-}X^* =$$

$$F^{-1}X^* = X.$$

$$\text{b) } XX^{*-}P_{X^*|V_{\varepsilon^*}Z}FV_{\varepsilon}Z = XX^{*-}P_{X^*|V_{\varepsilon^*}Z}FV_{\varepsilon}F'Z = XX^{*-}P_{X^*|V_{\varepsilon^*}Z}V_{\varepsilon^*}Z = 0.$$

y por lo tanto $XX^{*-}P_{X^*|V_{\varepsilon^*}Z}F$ es un operador de proyección oblicuo sobre $\mathcal{M}(X)$ a lo largo de $\mathcal{M}(V_{\varepsilon}Z)$.

■

www.bdigital.ula.ve

Capítulo 6

Aplicación de un modelo en dos etapas

A la fecha han sido diversos las publicaciones sobre el rendimiento estudiantil; y ello se debe a los cambios que ocurren en el proceso de enseñanza aprendizaje como consecuencia de las variaciones que acontecen en los sistemas educativos, que influyen de manera directa en el rendimiento académico a lo largo del tiempo, de allí lo interesante del tema para los investigadores. Desde hace tiempo en nuestras universidades se ha venido presentado el problema del bajo rendimiento estudiantil.

En su aspiración por explicar el rendimiento estudiantil, los investigadores han utilizado métodos y técnicas estadísticas diversas. González (1988), utilizó el análisis de componentes principales y el análisis de correspondencias múltiples en su estudio sobre el rendimiento estudiantil, y entre sus resultados halló que el bajo rendimiento estudiantil era uno de los grandes problemas de nuestras universidades, y que las notas obtenidas por los bachilleres influyen de manera positiva sobre el rendimiento universitario, y sugieren la implantación de un sistema que fortalezca los conocimientos adquiridos en el bachillerato con el fin de mejorar el rendimiento estudiantil en la universidad. Lamentablemente, después de más de 20 años el bajo rendimiento sigue siendo uno de los problemas que afectan la vida de nuestras universidades, lo cual explica aún más el porqué siempre habrá investigadores interesados en estudiar el rendimiento estudiantil y los factores que sobre él influyen.

González (1989) y Garnica et al. (1991), utilizaron modelos LISREL, y Análisis Discriminante respectivamente, para estudiar el rendimiento estudiantil. González destaca que el rendimiento estudiantil es influenciado fundamentalmente por un factor aptitudinal, el cual a su vez es influenciado por la preparación del estudiante al momento de ingresar a la universidad y por su situación socioeconómica. Por su parte, Garnica et al. en su estudio encontraron que el rendimiento en la universidad está ligado a las destrezas numéricas del estudiante, y al promedio de bachillerato, y recomiendan la implantación de cursos de nivelación universitaria. Posteriormente Garnica (1997), en su estudio del rendimiento estudiantil, haciendo uso del Análisis de Componentes Principales en conjunto con el Análisis de Varianza (ANOVA) encontró que en planteles públicos y privados el deterioro en la calidad de la educación era evidente, siendo aun más notable en los institutos de educación públicos; y que apenas una minoría de los alumnos que ingresan a las universidades poseían una buena preparación integral.

Más recientemente, Valera et al. (2009), Ponsot et al. (2009), y Varela et al. (2009), utilizan modelos de regresión logística en sus estudios relacionados con el rendimiento estudiantil universitario. Valera et al., al estudiar las calificaciones de los primeros dos semestres, de los estudiantes de la Facultad de Ingeniería, de la Universidad de Los Andes, determinaron que el promedio de bachillerato tiene un efecto significativo sobre el rendimiento académico en el primer semestre. Paralelamente Ponsot et al. y Varela et al.; estudiaron las calificaciones de los estudiantes de la Facultad de Ciencias Económicas y Sociales, de la Universidad de Los Andes. Ponsot et al., encontraron que aquellos estudiantes que fueron buenos alumnos en secundaria tenían casi el doble de posibilidades de obtener una eficiencia elevada en sus estudios universitarios, en relación con los alumnos con bajas calificaciones en secundaria, y Varela et al., hallaron que el período académico que mejor pronostica el rendimiento estudiantil dentro de cada una de las carreras era el primero y que aquellos estudiantes con rendimientos mayores a 15 puntos en el primer período académico poseen mayores posibilidades de obtener un rendimiento bueno en la carrera.

En este orden de ideas, este capítulo pretende mostrar cómo un modelo en dos etapas regular se puede emplear para predecir la calificación obtenida en el Examen de Admi-

sión de la Universidad Nacional Experimental del Táchira (U.N.E.T.) por un aspirante a ingresar a esta casa de estudios, a partir de su calificación en el curso propedéutico dictado por la misma universidad, y cuyo propósito es fortalecer los conocimientos obtenidos en el bachillerato.

6.1. El problema del Rendimiento Estudiantil

El rendimiento estudiantil siempre ha sido un tema de interés para muchos investigadores y para su estudio se han utilizado diferentes metodologías y diversas técnicas estadísticas. Los conocimientos adquiridos en el bachillerato por los estudiantes que aspiran ingresar al subsistema de educación universitaria son muy deficientes, y ello ha sido de gran preocupación para las universidades venezolanas en general, ya que constituye uno de los factores principales del bajo rendimiento académico. En el caso específico de la Universidad Nacional Experimental del Táchira (U.N.E.T., 2011) para buscar solventar esta situación se fundó la Unidad de Admisión, la cual "tiene como objetivo primordial, garantizar el ingreso de aspirantes a cursar estudios en las diferentes carreras que ofrece la universidad, de acuerdo a lo estipulado en las normas", y es el ente de la universidad encargado de todo lo relacionado con el curso propedéutico, el cual se dicta todos los semestres; y según la U.N.E.T. (2011), dicho curso cubre un conjunto de objetivos agrupados por áreas y tiene la finalidad de nivelar los conocimientos básicos del bachiller, preparándolo para que enfrente con menor dificultad las exigencias académicas del nivel básico de la universidad o del examen de admisión"

En este capítulo se han planteado dos objetivos fundamentales: fomentar el uso de los modelos lineales en dos etapas en las diferentes áreas del conocimiento, ya que hasta la fecha su uso por parte de los investigadores ha sido restringido a pesar del gran potencial que este tipo de modelos puede tener; y estudiar un problema en el campo de la educación superior al explicar la calificación obtenida por un bachiller en el examen de admisión de la U.N.E.T. en función de su calificación obtenida en el curso propedéutico de la U.N.E.T. a través de la aplicación de un modelo de dos etapas regular.

Para alcanzar estos objetivos, se utilizarán las expresiones de los estimadores del

vector de parámetros y del vector de medias del modelo en dos etapas regular desarrolladas en el capítulo 4 de este trabajo, resultados que ya han sido publicados en Sinha et. al. (2013); en las cuales estos estimadores se presentan como funciones de operadores de proyección.

6.1.1. Aplicación del modelo en dos Etapas Regular

Sea el modelo en dos etapas regular definido en (4.13), el cual por conveniencia reescribimos a continuación.

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ D & X_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}$$

$$y = X\beta + \varepsilon$$

con

$$Cov(\varepsilon) = \begin{bmatrix} \sigma_1^2 V_1 & 0 \\ 0 & \sigma_2^2 V_2 \end{bmatrix} = V_\varepsilon,$$

donde X_1, X_2 son de rango completo por columnas, V_1, V_2 son definidas positivas, y $\sigma_1^2 > 0, \sigma_2^2 > 0$ parámetros de covarianza desconocidos.

Obsérvese que lo que hace interesante a este tipo de modelos es que permite incorporar datos disponibles de algún estudio previo, siempre y cuando la variable respuesta en cada etapa sea la misma. Esto es posible gracias al término $D\beta_1$, que representa el término que enlaza las dos etapas del modelo.

De acuerdo a las expresiones dadas en el inciso 3 del Teorema 32, los UBLUE de β y $X\beta$ para el modelo (4.13) para todo $\sigma_1^2 > 0, \sigma_2^2 > 0$ están dados por:

$$\hat{\beta} = X^{*-} R_{X^*|V_\varepsilon, Z} F y \quad (6.1)$$

$$\widehat{X\beta} = X X^{*-} R_{X^*|V_\varepsilon, Z} F y \quad (6.2)$$

donde

$$R_{X^*|V_2^*Z} = \begin{bmatrix} I_1 - P'_{Z_1V_1} & 0 \\ DQ P'_{Z_1V_1} & I_2 - P'_{Z_2V_2} \end{bmatrix}$$

$$X^{*-} = \begin{bmatrix} X_1^- & 0 \\ 0 & X_2^- \end{bmatrix} = \begin{bmatrix} (X_1'X_1)^{-1} X_1' & 0' \\ 0 & (X_2'X_2)^{-1} X_2' \end{bmatrix}$$

y las matrices $P_{Z_1V_1}$, $P_{Z_2V_2}$ son los operadores de proyección definidos por Rao (1974), dadas por $P_{Z_1V_1} = Z_1\{Z_1V_1Z_1\}^-Z_1V_1$, $P_{Z_2V_2} = Z_2\{Z_2V_2Z_2\}^-Z_2V_2$ con $Z_1 = I_1 - P_{X_1}$, $Z_2 = I_2 - P_{X_2}$, y P_{X_i} la matriz operador proyección perpendicular sobre $\mathcal{M}(X_i)$ para $i = 1, 2$, siendo $\{Z_1V_1Z_1\}^-$, $\{Z_2V_2Z_2\}^-$ inversas generalizadas en el sentido de Rao (1973, p.24).

6.1.2. Población y muestra

Los datos utilizados para esta aplicación del modelo en dos etapas fueron obtenidos de la página web "http://www.unet.edu.ve/la-docencia/admision.html", de la Universidad Nacional Experimental del Táchira (U.N.E.T.) en la sección resultados, en el 2011. De ella se obtuvieron dos conjuntos de datos. El primero correspondiente al período 2009-1 con 661 registros y el segundo al período 2010-1 con 603 registros. Cada conjunto de datos contiene solamente los resultados del examen de admisión y las notas del curso propedéutico para cada uno de los aspirantes a cursar estudios en las carreras de Ingeniería Industrial, Mecánica, Electrónica, Informática y Civil. Se excluyeron los datos de aquellos bachilleres que no tienen nota en el Examen de Admisión o en el Curso Propedéutico y cuyo ingreso fue por admisión directa o a través de la Oficina de Planificación del Sector Universitario (OPSU), de la República Bolivariana de Venezuela.

En este estudio se asume que existe independencia entre los conjuntos de datos de los períodos 2009-1, 2010-1 debido a que pertenecen a grupos diferentes de estudiantes, quienes realizaron el curso propedéutico y presentaron el examen de admisión en períodos de tiempo diferentes no consecutivos.

Para el ajuste del modelo se obtuvo de cada período muestras aleatorias en lugar

de utilizar la totalidad de los datos disponibles. Dos razones motivan esta decisión, la primera está relacionada con el tipo de datos; cuando las variables son del tipo socioeconómico, los datos en general son bastante heterogéneos y los modelos ajustados a partir de la totalidad de los datos resultan usualmente poco precisos. En estos casos, los resultados basados en muestras más homogéneas que el conjunto de datos en su totalidad son mucho mejores. La segunda razón es que en muchas de las investigaciones que se realizan es difícil disponer de gran cantidad de datos por diversas razones, entre ellas económicas, o relacionadas con la dificultad de obtener las muestras. En este sentido, se desea mostrar que no necesariamente se requiere disponer de gran cantidad de datos para realizar el ajuste de un modelo en dos etapas regular.

El modelo en dos etapas regular requiere independencia de los datos de una etapa, con respecto a los datos de otra etapa y la variable respuesta a estudiar debe ser la misma en cada una de las etapas. En este caso, de cada conjunto de datos extraeremos muestras aleatorias independientes. En este sentido, del período 2009-1 se obtendrán muestras aleatorias de las variables Calificación en el Examen de Admisión Período 2009-1 ($CEA2009_1$) y Calificación en el Curso Propedéutico Período 2009-1 ($CP2009_1$), y del período 2010-1 se obtendrán las variables Calificación en el Examen de Admisión Período 2010-1 ($CEA2010_1$) y Calificación en el Curso Propedéutico Período 2010-1 ($CP2010_1$).

En relación con el modelo en dos etapas regular, las variables dependientes son $y_1 = CEA2009_1$ y $y_2 = CEA2010_2$. Las matrices X_1 y X_2 están formadas por dos columnas de datos. La primera es una columna de unos que denotamos con J , correspondiente al término de la pendiente de la etapa y la segunda es la variable explicativa. En este caso $CP2009_1$, $CP2010_1$ son las variables explicativas de la etapa 1 y la etapa 2 respectivamente.

Claramente $y_1 = CEA2009_1$, $X_1 = [J, CP2009_1]$ se corresponden con el período 2009-1 y $y_2 = CEA2010_2$, $X_2 = [J, CP2010_1]$ con el período 2010-1. En este estudio particular se decidió que el tamaño de cada una de las muestras utilizadas fuese mayor a 30 con el fin de garantizar la convergencia de las estimaciones hacia el verdadero valor de los parámetros y obtener buenos resultados en el caso de que el comportamiento de

los datos se aleje del comportamiento normal. En particular se tomó $n_1 = n_2 = n = 40$.

Es importante destacar que para la obtención de las muestras aleatorias, el ajuste del modelo y los resultados de este artículo, se utilizó un lenguaje de cómputos matemáticos y software estadístico. En la actualidad existen en el mercado paquetes especializados como MATLAB, Mathematica, Maple, Statgraphics, SPSS, etc. que se pueden utilizar para obtener estos resultados.

6.2. Ajuste del modelo en dos etapas

6.2.1. Análisis descriptivo y ajuste de los datos a una distribución normal

De cada período se obtuvo una muestra aleatoria de tamaño 40. El Cuadro 6.1 muestra el resultado de las correlaciones entre las variables $CEA2009_1$, $CP2009_1$, $CEA2010_1$, y $CP2010_1$ para la muestra seleccionada. En este se destaca el valor p de significancia.

	$CEA2009_1$	$CP2009_1$	$CEA2010_1$	$CP2010_1$
$CEA2009_1$	1	0,7743 0,0000*	-0,1287 0,4288	-0,1136 0,4853
$CP2009_1$	0,7743 0,0000*	1	0,1099 0,4995	0,1089 0,5035
$CEA2010_1$	-0,1287 0,4288	0,1099 0,4995	1	0,7590 0,0000*
$CP2010_1$	-0,1136 0,4853	0,1089 0,5035	0,7590 0,0000*	1

*Correlación estadísticamente significativa al 1 %

Fuente: Elaboración propia.

Cuadro 6.1: Correlaciones entre las variables bajo estudio.

En este cuadro también se observa que existen correlaciones estadísticamente significativas entre las variables de un mismo período (valores de $p = 0,0000$) y correlaciones no significativas entre variables de diferente período (valores de $p > 0,01$), lo que nos lleva a concluir que las muestras cumplen con el supuesto de independencia de los datos entre una etapa y otra, supuesto que requiere el Modelo en dos etapas regular. La Figura 6.1 permite apreciar la correlación existente entre los datos de las muestras. Obsérvese

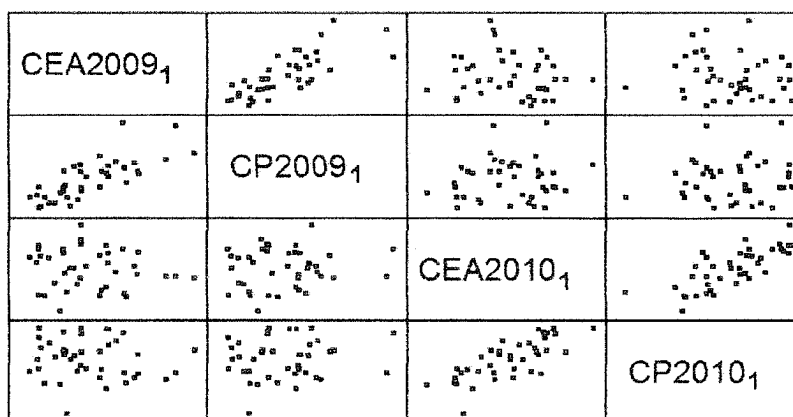


Figura 6.1: Fuente: Elaboración propia. Correlaciones entre las variables de la muestra

que sólo entre las variables para un mismo período se aprecia que existe una tendencia lineal entre las variables, en los demás casos se aprecia una nube de puntos sin patrón aparente en su forma.

En el Cuadro 6.2, se presentan algunas estadísticas descriptivas de interés. En éste se aprecia que las calificaciones obtenidas por los bachilleres en ambos períodos no son altas, en general. En ambos períodos, las calificaciones promedio son inferiores a 50 pts. Menos de la mitad en la escala del 0 a 100 puntos. El segundo cuartil más grande es 53,63 puntos, lo cual indica que 75 % de los aspirantes a ingresar a la U.N.E.T. en los períodos 2009-1, 2010-1, para esta muestra obtuvieron calificaciones inferiores a 54 puntos tanto en el curso propedéutico como en el examen de admisión. Esto refuerza la preocupación de docentes e instituciones en relación con la poca calidad de la educación con que llegan los bachilleres a las casas de estudios universitarios. Obsérvese que para las variables $CP2009_1$ y $CP2010_1$, que son las calificaciones alcanzadas por los alumnos de la muestra en los cursos propedéuticos de la U.N.E.T., la calificación más alta de la muestra en ambos períodos es 43,00 puntos.

En lo que respecta a la distribución de las variables calificación en el examen de admisión período 2009-1, y calificación en el examen de admisión período 2010-1, según los resultados obtenidos del ajuste de los datos de la muestra, dados en el Cuadro 6.3, estas presentan un comportamiento normal.

	CEA2009 ₁	CP2009 ₁	CEA2010 ₁	CP2010 ₁
Frecuencia	40	40	40	40
Media	49,95	15,52	41,25	30,08
Mediana	47,99	15,34	41,65	31,30
Desviación típica	7,10	8,96	7,04	8,30
Mínimo	39,00	2,75	25,83	7,00
Máximo	68,03	40,75	56,75	43,00
Rango	29,03	38,00	30,92	36,00
Primer cuartil	44,64	8,17	35,66	24,42
Tercer cuartil	53,63	21,00	46,95	36,09
Coef. de variación	14,36 %	57,71 %	17,08 %	27,59 %

Fuente: Elaboración propia.

Cuadro 6.2: Resumen estadístico de las variables.

6.3. Ajuste del modelo para los datos de la muestra

A partir de las ecuaciones (6.1) y (6.2) se procede a realizar el ajuste del Modelo en dos etapas. Para el ajuste del modelo se considerará que las matrices V_1 , V_2 poseen la forma de simetría compuesta. En otras palabras, se asumirá que los elementos seleccionados en cada una de las muestras están igualmente correlacionados. De esta manera, las matrices V_1 , V_2 son de la forma

$$V_i = \begin{bmatrix} 1 & \rho_i & \cdots & \rho_i \\ \rho_i & 1 & \cdots & \rho_i \\ \vdots & \vdots & \ddots & \vdots \\ \rho_i & \rho_i & \cdots & 1 \end{bmatrix}$$

para $i = 1, 2$ y $\rho_1 = 0,3$, $\rho_2 = 0,3$.

Variable: CEA2009 ₁				Variable: CEA2010 ₁			
Estadístico EDF	Valor	Forma Modificada	Valor P	Estadístico EDF	Valor	Forma Modificada	Valor P
Kolmogorov-Smirnov D	0,1240	0,7996	$\geq 0,10$	Kolmogorov-Smirnov D	0,0907	0,5852	$\geq 0,10$
Anderson-Darling Λ^2	0,5269	0,5375	0,1685	Anderson-Darling Λ^2	0,4000	0,4081	0,3470

Fuente: Elaboración propia.

Cuadro 6.3: Ajuste de las variables respuesta a una distribución normal.

En relación con la matriz D , debe seleccionarse de manera que satisfaga la condición $\mathcal{M}(D) \subset \mathcal{M}(X_2)$ señalada en el Teorema 32. Debido a que en el Modelo en dos

etapas $D\beta_1$ representa el término de enlace entre las etapas, el objetivo es proponer una matriz D en la que la influencia de la etapa 1 se refleje en la etapa 2. De la condición anterior, podemos escribir $D = X_2A$, donde A es una matriz de dimensión $p_2 \times p_1$. En consecuencia, se sigue que la segunda etapa puede expresarse como

$$\begin{aligned} y_2 &= X_2A\beta_1 + X_2\beta_2 + \varepsilon_2 \\ &= X_2(A\beta_1 + \beta_2) + \varepsilon_2 \\ &= X_2\beta_2^* + \varepsilon_2 \end{aligned}$$

De esta manera se aprecia que la influencia de la etapa 1 en la etapa 2 está relacionada a la estructura de la matriz A . En el caso particular en que A sea

$$A = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \quad (6.3)$$

$$\beta_2^* = A\beta_1 + \beta_2 = \begin{bmatrix} \beta_{21} \\ \beta_{12} + \beta_{22} \end{bmatrix}$$

luego, la segunda etapa puede escribirse como

$$\begin{aligned} y_2 &= \beta_{21}J + (\beta_{12} + \beta_{22})X_{22} + \varepsilon_2 \\ &= \beta_{21}^*J + \beta_{22}^*X_{22} + \varepsilon_2 \end{aligned}$$

donde recordemos J es un vector de unos de $n_2 \times 1$ y X_{22} la variable independiente de la etapa 2; esto es, $X_{22} = CP2010_1$. Otras opciones para la matriz A son

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

En particular, estudiaremos el caso donde la matriz A esta dada por (6.3). Los resultados obtenidos a partir de la muestra seleccionada se señalan en el Cuadro 6.4.

En este cuadro vemos que $\widehat{\beta}_{21} = \widehat{\beta}_{21}^* = 21,88$ y $\widehat{\beta}_{22}^* = \widehat{\beta}_{12} + \widehat{\beta}_{22} = 0,64$. Recuérdese

Coeficientes de la Etapa 1			Coeficientes de la Etapa 2			
	$\widehat{\beta}_{11}$	$\widehat{\beta}_{12}$	$\widehat{\beta}_{21}$	$\widehat{\beta}_{22}$	$\widehat{\beta}_{21}^*$	$\widehat{\beta}_{22}^*$
$D = X_2A$	39,929	0,61373	21,876	0,030331	21,876	0,64406

Cuadro 6.4: Resultados de ajustar el modelo en dos etapas.

que las variables dependientes son las calificaciones en los exámenes de admisión en cada período, y las variables independientes las calificaciones en los cursos propedéuticos.

La calificación obtenida por un estudiante en el curso propedéutico depende de muchos factores, los cuales se han dividido en dos grupos. Los factores inherentes al estudiante como lo son su ambiente familiar, su condición económica, y sobre todo los conocimientos adquiridos en el bachillerato, entre otros, y los factores relacionados con el curso propedéutico entre los que destacan la orientación profesoral con base en un material educativo seleccionado para reforzar sus conocimientos del bachillerato, la interrelación con otros compañeros en el curso propedéutico que tienen como meta común ingresar al sistema de educación superior, lo cual fomenta en el estudiante una sana competencia por adquirir conocimientos, así como otros factores relacionados.

En este sentido, podemos interpretar la influencia de la etapa 1 sobre la 2, representada por el coeficiente $\widehat{\beta}_{12} = 0,614$, como la influencia de los factores relacionados con el curso propedéutico, la cual cambia en la medida que se lleven a cabo correctivos que mejoran la eficiencia de dicho curso. De esta manera, la influencia de la calificación del curso propedéutico del período 2010-1 ($CP2010_1$) sobre la calificación en el examen de admisión para el mismo período ($CEA2010_1$), puede expresarse como la suma de dos componentes. El componente relacionado con los factores propios del curso propedéutico $0,614 * CP2010_1$ y el componente correspondiente a los factores inherentes al estudiante $0,03 * CP2010_1$. La influencia total de la calificación en el curso propedéutico del período 2010-1 sobre su calificación en el examen de admisión está dada por $0,644 * CP2010_1$. Esta forma de interpretar los coeficientes refleja la realidad en lo que respecta a la mala preparación con la que egresan la mayoría de los estudiantes de hoy día de los liceos del país. De esta manera, el modelo correspondiente a la etapa 2 se expresa como:

$$CEA2010_{1i} = 21,88 + 0,644CP2010_{1i} + \varepsilon_{2i}, i = 1, 2, \dots, 40.$$

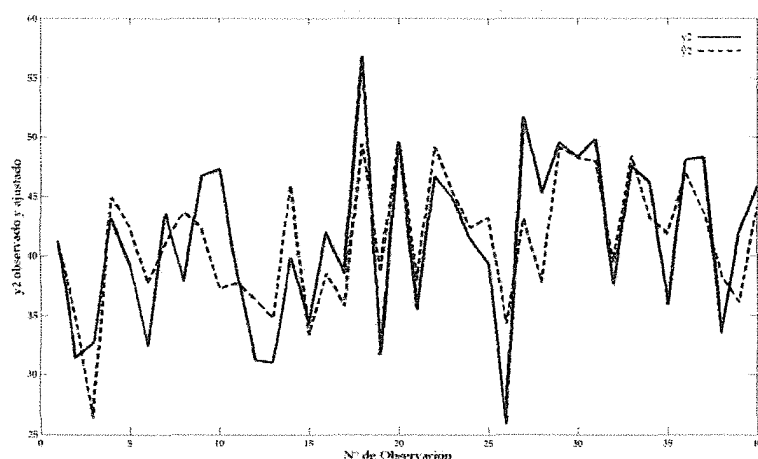


Figura 6.2: Valores observados y predichos a partir del modelo ajustado. Fuente: Elaboración propia.

Gr	β_{11}			β_{12}			β_{21}			β_{22}			β_{22}^*		
	Prom	Min	Max	Prom	Min	Max	Prom	Min	Max	Prom	Min	Max	Prom	Min	Max
1	38,75	35	41,26	0,62	0,52	0,81	22,17	13,09	25,78	0,01	-0,19	0,32	0,64	0,5	0,94
2	39,82	37,17	43,17	0,6	0,44	0,72	24,62	19,78	26,96	-0,03	-0,11	0,15	0,57	0,48	0,73
3	40,57	39,25	42,5	0,55	0,44	0,66	21,23	11,97	26,13	0,11	-0,08	0,35	0,67	0,51	0,94
4	40,76	36,85	42,93	0,56	0,45	0,77	22,46	16,58	25,24	0,06	-0,19	0,24	0,62	0,55	0,81
5	39,24	35,82	42,58	0,6	0,49	0,73	22,23	14,05	26,09	0,03	-0,21	0,29	0,63	0,52	0,89
6	39,95	38,15	41,47	0,57	0,45	0,67	21,41	16,98	24,7	0,1	-0,03	0,26	0,67	0,59	0,79
7	40,28	37,97	42,01	0,58	0,52	0,65	23,4	19,31	26,45	0,01	-0,09	0,12	0,59	0,52	0,71
8	39,48	38,35	41,44	0,58	0,51	0,65	21,71	15,66	25,29	0,06	-0,07	0,28	0,64	0,56	0,82
9	39,49	38,47	40,32	0,6	0,55	0,66	21,37	18,62	24	0,05	-0,05	0,19	0,65	0,57	0,74
10	40,13	38,85	41,78	0,57	0,5	0,62	21,88	18,81	25,05	0,08	-0,01	0,27	0,65	0,54	0,77
11	39,58	38,1	41,34	0,58	0,51	0,69	23,22	20,47	26,44	0,04	-0,12	0,15	0,62	0,53	0,71
12	39,47	36,79	41,21	0,6	0,52	0,74	21,59	17,11	24,91	0,04	-0,11	0,23	0,64	0,55	0,76
13	39,08	35,97	41,27	0,61	0,5	0,75	23,43	20,1	28,61	0	-0,16	0,15	0,61	0,47	0,7
14	38,81	35,08	40,84	0,63	0,55	0,78	21,94	16,37	26,25	0,02	-0,27	0,26	0,64	0,51	0,85
15	39,15	37,73	41,59	0,63	0,53	0,72	20,29	13,84	24,8	0,06	-0,03	0,25	0,69	0,56	0,88
16	39,11	35,99	41,7	0,62	0,52	0,76	22,11	16,55	25,7	0	-0,14	0,12	0,62	0,53	0,8
17	39,22	35,54	42,61	0,63	0,48	0,83	21,05	16,45	26,83	0,03	-0,08	0,22	0,66	0,5	0,82
18	39,04	37,14	41,62	0,63	0,52	0,77	21,35	16,42	26,65	0,03	-0,15	0,25	0,66	0,51	0,78
19	39,61	37,76	42,82	0,61	0,49	0,69	21,31	17,45	27,99	0,05	-0,18	0,24	0,66	0,48	0,79
20	38,68	34,65	40,96	0,64	0,52	0,79	21,37	18,97	25,63	0,01	-0,21	0,18	0,65	0,54	0,75

Fuente: Elaboración propia.

Cuadro 6.5: Promedios, mínimos y máximos de los parámetros por grupo.

Como una manera empírica de validar el modelo se obtuvo otra muestra aleatoria de tamaño 40 del período 2010-1, y se ajustó el modelo anterior obteniendo el vector de valores predichos $\widehat{CEA2010_1} = \hat{y}_2$ (vector de valores ajustados). Luego se procedió a obtener el vector de *error* $= y_2 - \hat{y}_2 = \widehat{CEA2010_1} - \widehat{CEA2010_1}$. Los resultados se muestran en las Figuras 6.2 y 6.3. En la 6.2, se aprecia que los valores predichos por el modelo (línea segmentada) son bastante cercanos a los observados (línea continua), lo cual se aprecia mejor en la Figura 6.3 donde ninguna diferencia entre los valores observados y los valores predichos de la $\widehat{CEA2010_1}$ obtenidos a partir del modelo es mayor a 10 puntos.

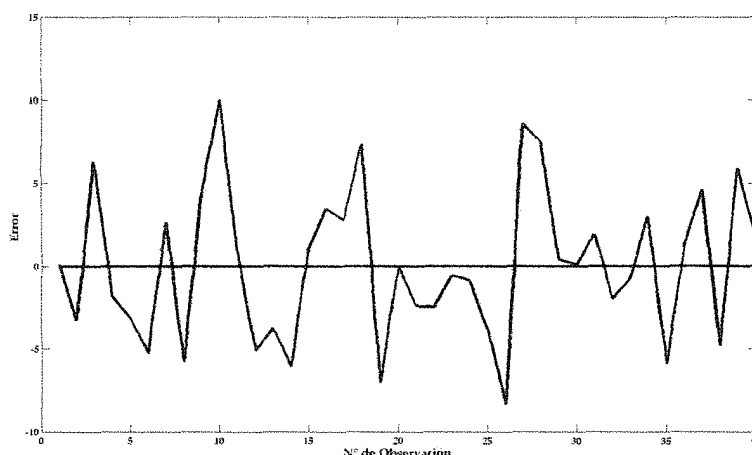


Figura 6.3: Error = Valores observados - Valores predichos. Fuente: Elaboración propia.

6.4. Ajuste del modelo para varias muestras

En la sección 6.2.1 se mostró que las variables $CEA2009_1$ y $CEA2010_1$ poseen un comportamiento normal, y de la ecuación (6.1) es claro que $\hat{\beta}$ es una función lineal de vector $y = (y'_1, y'_2)'$, por lo tanto el vector aleatorio $\hat{\beta}$ también se distribuye en forma normal. En consecuencia, $\hat{\beta}_1$, $\hat{\beta}_2$ y cada elemento $\hat{\beta}_{ij}$ $i, j = 1, 2$ individualmente se distribuye también en forma normal. Por otra parte, un resultado bastante conocido en estadística dado por Mood et al. (1974, p. 250) expresa que si $X_1, \dots, X_n$ es una muestra aleatoria proveniente de una distribución normal con media μ y varianza σ^2 , la variable aleatoria

$$\begin{aligned} T &= \frac{(\bar{X} - \mu)/(\sigma/\sqrt{n})}{\sqrt{(1/\sigma^2) \sum_{i=1}^n (X_i - \bar{X})^2 / (n-1)}} = \frac{(\bar{X} - \mu)}{\sigma/\sqrt{n}} \frac{1}{\sqrt{(S^2/\sigma^2)}} \\ &= \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \frac{\sigma}{S} = \frac{\sqrt{n}(\bar{X} - \mu)}{S} \\ &= \frac{(\bar{X} - \mu)}{S/\sqrt{n}} \end{aligned}$$

tiene distribución t con $n - 1$ grados de libertad, donde S^2 es la varianza de la muestra dada por

$$S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n - 1}$$

el intervalo de confianza del $(1 - \alpha) \times 100\%$ para $E(X) = \mu$ en el caso cuando la varianza σ^2 sea desconocida, esta dado por

$$(\bar{X} - t_{\alpha/2}S/\sqrt{n}, \bar{X} + t_{\alpha/2}S/\sqrt{n}) \quad (6.4)$$

Gr	β_{11}		β_{12}		β_{21}		β_{22}		$\hat{\beta}_{22}$	
	L_i	L_s	L_i	L_s	L_i	L_s	L_i	L_s	L_i	L_s
1	37,293	40,197	0,562	0,679	19,591	24,748	-0,082	0,11	0,552	0,719
2	38,477	41,158	0,545	0,651	23,112	26,132	-0,092	0,028	0,513	0,619
3	39,868	41,267	0,508	0,6	18,329	24,128	0,018	0,208	0,576	0,758
4	39,534	41,987	0,501	0,625	20,679	24,251	-0,032	0,147	0,565	0,676
5	37,775	40,697	0,552	0,652	19,595	24,869	-0,081	0,137	0,542	0,717
6	39,063	40,83	0,518	0,626	19,858	22,967	0,023	0,175	0,63	0,713
7	39,413	41,138	0,54	0,612	21,775	25,026	-0,036	0,064	0,545	0,634
8	38,747	40,207	0,545	0,622	19,694	23,718	-0,013	0,127	0,585	0,696
9	38,983	39,999	0,575	0,621	20,172	22,574	0,01	0,094	0,613	0,686
10	39,345	40,906	0,543	0,6	20,219	23,533	0,018	0,136	0,597	0,7
11	38,816	40,345	0,532	0,632	21,744	24,687	-0,026	0,106	0,582	0,661
12	38,512	40,427	0,538	0,658	19,692	23,488	-0,056	0,132	0,579	0,694
13	37,623	40,531	0,542	0,672	21,437	25,424	-0,071	0,068	0,55	0,662
14	37,363	40,266	0,575	0,677	19,764	24,107	-0,093	0,124	0,568	0,715
15	38,297	40,007	0,595	0,666	17,717	22,855	-0,001	0,127	0,613	0,775
16	37,617	40,595	0,57	0,673	20,181	24,034	-0,063	0,064	0,566	0,679
17	37,236	41,211	0,542	0,71	18,526	23,565	-0,04	0,105	0,589	0,728
18	37,971	40,117	0,573	0,681	19,24	23,455	-0,061	0,117	0,589	0,722
19	38,63	40,588	0,567	0,651	19,04	23,578	-0,042	0,144	0,595	0,725
20	37,12	40,235	0,576	0,707	19,677	23,054	-0,063	0,089	0,601	0,707

Fuente: Elaboración propia.

Cuadro 6.6: Intervalos de confianza para los parámetros del modelo por grupo de muestras.

Tomando en consideración estos resultados se procedió a obtener 20 grupos de muestras, cada uno formado por 10 muestras aleatorias e independientes de tamaño 40. Luego para cada muestra, en cada uno de los grupos se obtuvo el valor de los parámetros $\beta_1 = (\beta_{11}, \beta_{12})$, $\beta_2 = (\beta_{21}, \beta_{22})$ y $\beta_2^* = (\beta_{21}^*, \beta_{22}^*)$. Posteriormente, se obtuvieron algunas estadísticas descriptivas e intervalos de confianza para cada uno de los parámetros, para cada uno de los grupos. Los promedios y los valores mínimos y máximos se dan en el Cuadro 6.5. Los intervalos de confianza se dan en el Cuadro 6.6. Al comparar los valores de $\hat{\beta}_{11}$, $\hat{\beta}_{12}$, $\hat{\beta}_{21} = \hat{\beta}_{21}^*$, $\hat{\beta}_{22} = \hat{\beta}_{22}^*$ dados en el Cuadro 6.4 con los valores del Cuadro 6.5 se

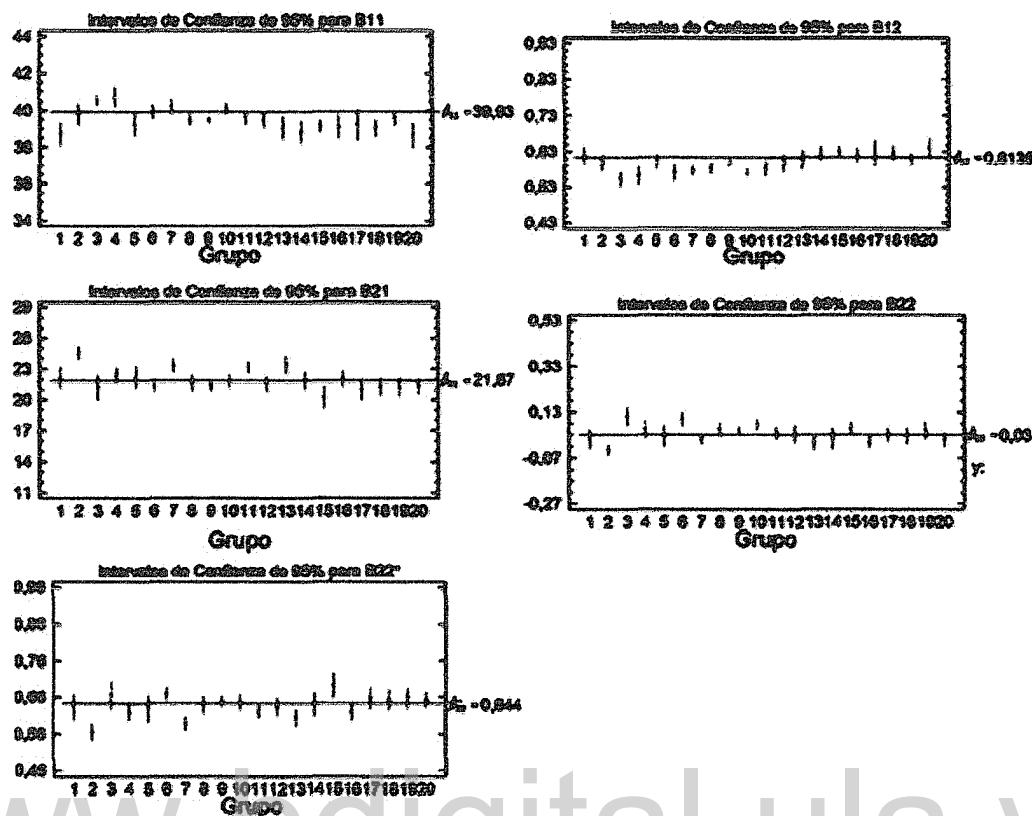


Figura 6.4: Intervalos de confianza para los parámetros del modelo en dos etapas. Fuente: Elaboración propia

observa que en general ellos se ubican entre los valores mínimo y máximo encontrados para cada grupo de muestras. Por ejemplo, $\hat{\beta}_{11} = 39,93$ se encuentra entre 35,00 y 41,26 que son los valores mínimo y máximo de los $\hat{\beta}_{11}$ del grupo 1. Esto muestra de manera empírica que los valores del Cuadro 6.4 utilizados para ajustar el modelo de la etapa 2 son similares en orden de magnitud a sus pares obtenidos en cada una de las 10 muestras, en cada uno de los 20 grupos. En el Cuadro 6.6 se muestran los intervalos de confianza para cada uno de los parámetros y para cada uno de los 20 grupos, obtenidos a partir de la expresión (6.4).

En la Figura 6.4 se visualizan los intervalos de confianza dados en el Cuadro 6.6 y una línea de referencia que señala el valor del parámetro correspondiente del Cuadro 6.4. En él se aprecia que los valores del Cuadro 6.4 se encuentran en muchos de los intervalos de confianza, lo cual valida estos resultados como representativos de los valores reales

de los parámetros que ellos estiman, mostrando que no necesariamente se debe disponer de gran cantidad de datos para ajustar un modelo si la muestra obtenida para ajustar los datos es lo suficiente representativa, de modo que las variables independientes sean buenas variables predictivas de la variable dependiente de interés a estimar.

6.5. Conclusiones

Al ajustar un modelo en dos etapas regular para predecir las calificaciones obtenidas por los bachilleres en el examen de admisión de la U.N.E.T. a partir de su calificación obtenida en el curso propedéutico previo a dicho examen, hemos mostrado la potencialidad que este tipo de modelado puede tener en otras áreas del saber. Las bondades del modelo de permitir incorporar información correspondiente a los períodos 2009-1 y 2010-1 en un solo modelo se hicieron evidentes al poder separar la influencia de la calificación obtenida en el propedéutico del 2010-1 en dos componentes. Un componente representado por lo que se ha denominado la influencia histórica del curso propedéutico, en donde se refleja la técnica, los materiales, métodos y sobre todo la experiencia adquirida en el curso propedéutico a lo largo de los años, como ente dedicado a la nivelación y al apuntalamiento del conocimiento adquirido por los jóvenes en el bachillerato, representada por el coeficiente β_{12} , y un segundo componente relacionado con los factores inherentes al estudiante señalados anteriormente, representados en el coeficiente β_{22} .

En este capítulo se ha mostrado que la implantación del modelo en dos etapas es un procedimiento relativamente sencillo. Especialmente resulta de utilidad cuando se le emplea para predecir la calificación en el examen de admisión 2010-1 a partir de muestras representativas (de tamaño 40 o superiores), seleccionadas en forma aleatoria e independiente de cada período escolar estudiado. Se espera que con este caso de aplicación, haber contribuido en fomentar el uso de los modelos en etapas desde la concepción de los operadores de proyección, en particular el uso de los modelos en dos etapas regular, como herramientas útiles para el análisis de datos. Finalmente se desea destacar que los resultados obtenidos en este capítulo ya fueron publicados por Valera et. al.(2011).

Bibliografía

Alalouf, I.S., y Styan, G.P.H. (1979). *Estimability and testability in restricted linear models*. Statistics, 10: 2, pp.189-201.

Baksalary, J. y Kala, R. (1981). *Linear Transformations Preserving Best Linear Unbiased Estimators in a General Gauss-Markoff Model*. The Annals of Statistics, Vol. 9, No. 4, pp. pp.913-916.

Bhimasankaran, P. y Sengupta, D. (1996). *The linear zero functions approach to linear models*. Sankhyā: The Indian Journal of Statistics, Series B Vol.58, No.3, pp.338-351.

Christensen, R. (2011). *Plane Answers to Complex Questions. The Theory of Linear Models*. Fourth Edition. Springer.

Christensen, R. y Lin Y. (2012). *Linear models that allow perfect estimation*. Statistical Papers.

Drygas, H. (1983). *Sufficiency and Completeness in the General Gauss-Markov Model*. Sankhyā: The Indian Journal of Statistics, Series A, Vol. 45, No. 1, pp.88-98.

Garnica, E. (1997). *El rendimiento estudiantil: una metodología para su medición*. Revista Economía. N° 13. Universidad de los Andes. pp.7-25.

Garnica, E., González P., Díaz, A. y Torres, E. (1991). *Análisis discriminante: estudio del rendimiento estudiantil*. Revista Economía. N° 6. Universidad de los Andes. pp.63-95.

González M., P. (1988). *Indicadores Sintéticos del Rendimiento Estudiantil*. Revista Economía. N° 2. Universidad de los Andes. pp.69-84.

González M., P. (1989). *Aplicación del Lisrel al Análisis del Rendimiento Estudiantil*. Revista Economía. N° 4. Universidad de los Andes. pp.73-96.

Graybill, F.A. (1976). *Theory and Application of the Linear Model*. Duxbury Press, North Scituate, MA.

GroB, J. (2004). *The general Gauss-Markov model with possibly singular dispersion matrix*. Statistical Papers. 45, pp.311-336

Gross, J., Trenkler, G. y Werner, H. (2001). *The Equality of Linear Transforms of the Ordinary Least Squares Estimator and the Best Linear Unbiased Estimator*. Sankhya: The Indian Journal of Statistics, Series A, Vol. 63, No. 1, pp. 118-127

Harville, D. (1997). *Matrix algebra from a statistician's perspective*. New York. Springer-Verlag.

Hernández, R., Fernández, C. y Baptista P. (1998). *Metodología de la Investigación*. (2ª. ed.). D. F., México: Mc Graw Hill.

Hernández, R., Fernández, C. y Baptista P. (2006). *Metodología de la Investigación*. (5ª. ed.). D. F., México: Mc Graw Hill.

Kubáček, L. (1986). *Multistage regression model*. Applications of Mathematics,

Vol. 31, No. 2, pp.89-96.

Kubáček, L. (1988). *Two-stage regression model*. Math. Slovaca 38, pp.383-393.

Kubáček, L. (1993). *Two stage linear model with constraints*. Mathematica Slovaca, Vol. 43, No. 5, pp.643-658.

Liu, A. (1996). *Estimation of the parameters in two linear models with only some of the parameters vectors identical*. Statistics & Probability Letters 29, pp.369-375.

Mood, A., Graybill, F. y Boes D. (1987). *Introduction to the theory of statistics*. 3ra edición. New York: McGraw-Hill.

Mueller, J. (1987). *Sufficiency and Completeness in the Linear Model*. Journal of Multivariate Analysis 21, pp.312-323.

Ponsot, E., Sinha, S., Varela, L. y Valera, J. (2009). *Un Modelo de Regresión Logística en los Estudios Universitarios: Caso FACES ULA*. Actualidad Contable FACES. Año 12. N° 18. Mérida. Venezuela. pp.81-102.

Puntanen, S. (1996). *Some matrix results related to a partitioned singular linear model*. Communications in Statistics - Theory and Methods, 25: 2, pp.269-279.

Puntanen, S. (1997). *Some further results related to reduced singular linear models*. Communications in Statistics - Theory and Methods, 26: 2, pp.375-385.

Rao, C.R. (1973). *Linear Statistical Inference and Its Applications*. New York. Wiley.

Rao, C.R. (1974). *Projectors, Generalized Inverses and the BLUE's*. Journal of the Royal Statistical Society. Series B (Methodological), Vol. 36, No. 3, pp.442-448.

Risquez, G. (1999). *Metodología de la Investigación I. Manual teórico práctico*. (1^a. ed.) Maracaibo, Venezuela: Editorial Universo de Venezuela C.A.

Sinha, S., Goitía, A. y Valera, J. (2013). *UBLUE FOR THE REGULAR TWO-STAGE LINEAR MODEL FROM THE PERSPECTIVE OF PROJECTION OPERATORS*. Electronic Journal of Applied Statistical Analysis. Vol. 6, N° 1, pp.97-109

Universidad Nacional Experimental del Táchira (2011) [Página Web en Línea]. Disponible: <http://www.unet.edu.ve/la-docencia/admision.html>. [Consulta: 2011, agosto 15]

Valera, J., Sinha, S., Varela, J. y Ponsot E. (2009). *Una explicación del rendimiento estudiantil universitario mediante modelos de regresión logística*. Visión Gerencial FACES. Año 8. N° 2. Mérida. Venezuela. pp.415-427.

Varela, J., Sinha, S., Ponsot, E. y Valera, J. (2009). *Valor pronóstico del k-ésimo período inicial sobre rendimiento de los estudiantes de la FACES-ULA*. Actualidad Contable FACES. Año 12. N° 19. Mérida. Venezuela. pp.133-146.

Valera, J., Sinha, S. y Goitía A. (2011). *Estimación de calificación del examen de admisión usando el modelo en dos etapas regular: Caso U.N.E.T*. Actualidad Contable FACES. Año 14. N° 23. Mérida. Venezuela. pp.134-153.

Volaufova, J. (1987). *Estimation of parameters of mean and variance in two-stage linear models*. Applications of Mathematics, Vol. 32, No. 1, pp.1-8.

Volaufova, J. (1988). *Note on the estimation of parameters of the mean and the variance in n-stage linear models*. Applications of Mathematics, Vol. 33, No. 1, pp.41-48.

Volaufova, J. (2004). *Some estimation problems in multistage*. Linear Algebra and its Applications, 388, pp.389-397.

Wulff, S. y Birkes, D. (2005). *Best unbiased estimation of fixed effects in mixed ANOVA models*. Journal of Statistical Planning and Inference 133, pp.305-313.

Zhang, B.; Liu, B. and Lu, C. (2004). *A Study of the Equivalence of the BLUEs between a Partitioned Singular Linear Model and Its Reduced Singular Linear Models*. Acta Mathematica Sinica, English Series June, Vol.20, No.3, pp.557-568.

www.bdigital.ula.ve