Benemérita Universidad Autónoma de Puebla

FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS
LICENCIATURA EN MATEMÁTICAS

ESTIMACIÓN DE PARÁMETROS Y OPERADORES

TESIS

QUE PARA OBTENER EL TÍTULO DE
LICENCIADO EN MATEMÁTICAS

PRESENTA
MARIA FERNANDA PICHARDO ZAMORA

DIRECTOR DE TESIS
DR. FERNANDO VELASCO LUNA

PUEBLA, PUE. JULIO 2016

A Luis Pichardo Rodriguez
y Rafael Zamora Sanchez

Agradecimientos

En primer lugar quiero agradecer a quienes desde pequeña me formaron y supieron guiarme siempre por el camino correcto, con su ejemplo y su dedicación. A mi papá por inspirarme a ser un poquito de lo inteligente que es él, a mi mamá por mostrarme como se debe luchar y salir adelante en la vida. Gracias por darme siempre las bases para ser mejor persona, Agustin y Margarita.

Quiero darle gracias a mis hermanas, a la mayor por ser una guía, y a la menor por permitirnos ser un ejemplo para ella. Gracias por todo, ustedes también son mi motor, Lau y Ale.

Sin duda, el camino siempre es más pesado y largo si se recorre solo, yo tuve la fortuna de tener excelentes compañeros y más que compañeros amigos que estuvieron conmigo y me ayudaron en este trayecto. Gracias a todos y cada uno de ustedes.

Para el Dr. Fernando Velasco Luna mis más sinceros agradecimientos por el apoyo recibido en la realización de esta tesis, por las enseñanzas aprendidas, por el tiempo dedicado, porque siempre tuve su comprensión y apoyo.

Agradezco también a mi jurado por el tiempo dedicado en la revisión de este proyecto y por su colaboración.

Se agradece el apoyo otorgado a PROMEP 103.5/13/6823 para la realización de este trabajo.

Por último quiero agradecer a las personas que me impulsaron a terminar este proyecto y que me ayudan a que día a día crezca un poquito más.

Introducción

El álgebra es una de las áreas de la matemática de mayor importancia tanto desde el punto de vista teórico como aplicado. Una de las áreas de mayor aplicación de las matemáticas en general, y en particular del álgebra lineal es la estadística. La teoría de espacios vectoriales de dimensión finita proporciona un marco para trabajar conceptos de inferencia en el modelo lineal general (MLG). Conceptos como subespacio columna, operador proyector ortogonal u oblicuo, matriz inversa generalizada, entre otros, juegan un papel de suma importancia en el estudio del MLG. Por otro lado, los modelos lineales estadísticos juegan un papel muy importante en la teoría estadística. La teoría de los modelos lineales proporciona la base teórica para una gran variedad de técnicas estadísticas tales como análisis de regresión, análisis de varianza, análisis de covarianza, diseños experimentales, análisis discriminante, modelos multivariados estadísticos, entre otras. Estudiar la relación de una variable con otra es una actividad científica importante en casi toda rama de la ciencia. Es de importancia ya que nos puede capacitar para predecir el valor de una variable a partir del conocimiento de otras y esto puede ser útil para tareas futuras.

La caracterización de los estimadores de los coeficientes de regresión β y de la varianza σ^2 en el MLG por medio del operador proyector ortogonal $\mathbf{P_X} = \mathbf{X}\left(\mathbf{X}^t\mathbf{X}\right)^{-}\mathbf{X}^t$ u oblicuo $\mathbf{P_{XV}} = \mathbf{X}\left(\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}^t\mathbf{V}^{-1}$, permite la mejor comprensión de las propiedades, ya que está basada en los principios del operador proyector y del subespacio generado por las columnas de las matrices involucradas. En el criterio de mínimos cuadrados para definir el mejor ajuste, el estimador de mínimos cuadrados ordinarios del vector de coeficientes en el MLG se puede expresar en términos del operador proyector ortogonal $\mathbf{P_X}$ sobre el espacio columna $S(\mathbf{X})$; así mismo, el estimador de mínimos cuadrados generalizados se puede expresar en términos del operador proyector oblicuo $\mathbf{P_{XV}}$.

Nuestro objetivo es dar la caracterización de los estimadores de los coeficientes de regresión y de la varianza en el modelo lineal general en términos de los operadores proyectores. En particular, expresar la estimación de los coeficientes de regresión en términos de los operadoradores proyector, expresar la estimación de los coeficientes de la varianza en términos del operador proyector ortogonal, expresar la estimación de los coefientes de regresión en términos del operador proyector oblicuo y expresar la estimación de los coeficientes de la varianza en términos del operador proyector oblicuo.

Índice general

Capítulo 1

SUBESPACIO SUMA DIRECTA

En este capítulo se introduce el concepto de espacio vectorial, así como también algunas de sus propiedades, esto nos permitirá posteriormente introducirnos en el concepto de subespacio suma directa de subespacios, así como también algunos resultados de tal subespacio, los cuales se usan a lo largo del trabajo.

1.1. Espacio vectorial

En esta sección definiremos que es un espacio vectorial, para después introducir el concepto de combinación lineal, el cual será parte importante en el desarrollo de este trabajo.

Definición 1.1.1 *Sea $\mathbb{R}$ el conjunto de los números reales. Un* ***espacio vectorial*** *sobre $\mathbb{R}$ es un conjunto $V \neq \emptyset$ sobre el que hay definidas dos operaciones:*

1. *Adición:*

$$+ : V \times V \to V$$
$$(\boldsymbol{\alpha}, \boldsymbol{\beta}) \to \boldsymbol{\alpha} + \boldsymbol{\beta}$$

2. *Producto por un escalar:*

$$\cdot : \mathbb{R} \times V \to V$$

$$(c, \boldsymbol{\alpha}) \to c \cdot \boldsymbol{\alpha}$$

tales que cumplen las siguientes condiciones:

Respecto a la adición:

1. *Conmutatividad:* $\boldsymbol{\alpha} + \boldsymbol{\beta} = \boldsymbol{\beta} + \boldsymbol{\alpha}$, $\forall\ \boldsymbol{\alpha}, \boldsymbol{\beta} \in V$.
2. *Asociatividad:* $(\boldsymbol{\alpha} + \boldsymbol{\beta}) + \boldsymbol{\gamma} = \boldsymbol{\alpha} + (\boldsymbol{\beta} + \boldsymbol{\gamma})$, $\forall\ \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma} \in V$.
3. *Existencia del elemento neutro aditivo: Existe un único elemento* $\boldsymbol{0} \in V$, *tal que* $\boldsymbol{\alpha} + \boldsymbol{0} = \boldsymbol{0} + \boldsymbol{\alpha} = \boldsymbol{\alpha}$, $\forall\ \boldsymbol{\alpha} \in V$.
4. *Existencia del elemento inverso aditivo: Para cada* $\boldsymbol{\alpha} \in V$ *existe un único elemento* $-\boldsymbol{\alpha} \in V$ *tal que* $\boldsymbol{\alpha} + (-\boldsymbol{\alpha}) = (-\boldsymbol{\alpha}) + \boldsymbol{\alpha} = \boldsymbol{0}$.

Respecto al producto por un escalar:

5. $1 \cdot \boldsymbol{\alpha} = \boldsymbol{\alpha}$, $\forall\ \boldsymbol{\alpha} \in V$.
6. $c \cdot (b \cdot \boldsymbol{\alpha}) = (c \cdot b) \cdot \boldsymbol{\alpha}$, $\forall\ b, c \in \mathbb{R}$, $\forall\ \boldsymbol{\alpha} \in V$.
7. $(c + b) \cdot \boldsymbol{\alpha} = c \cdot \boldsymbol{\alpha} + b \cdot \boldsymbol{\alpha}\ \forall\ c, b \in \mathbb{R}$, $\forall\ \boldsymbol{\alpha} \in V$.
8. $c \cdot (\boldsymbol{\alpha} + \boldsymbol{\beta}) = c \cdot \boldsymbol{\alpha} + c \cdot \boldsymbol{\beta}\ \forall\ c \in \mathbb{R}$, $\forall\ \boldsymbol{\alpha}, \boldsymbol{\beta} \in V$.

Los elementos de un espacio vectorial se denominan **vectores**.

Proposición 1.1.1 *Sea* V *un espacio vectorial. Entonces, se tiene que*

1. $c \cdot \boldsymbol{0} = \boldsymbol{0}$, *para todo* $c \in \mathbb{R}$
2. $0 \cdot \boldsymbol{\alpha} = \boldsymbol{0}$, *para todo* $\boldsymbol{\alpha} \in V$.
3. *Si* $c \cdot \boldsymbol{\alpha} = \boldsymbol{0}$, *para* $c \in \mathbb{R}$ *y* $\boldsymbol{\alpha} \in V$ *entonces,* $c = 0$, *o bien* $\boldsymbol{\alpha} = \boldsymbol{0}$
4. $(-1) \cdot \boldsymbol{\alpha} = -\boldsymbol{\alpha}$, *para todo* $\boldsymbol{\alpha} \in V$.

Definición 1.1.2 *Un vector* $\boldsymbol{\beta}$ *de* V *se dice una* ***combinación lineal*** *de los vectores* $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$ *en* V, *si existen escalares* $c_1, c_2, \ldots, c_n$ *de* $\mathbb{R}$ *tales que*

$$\boldsymbol{\beta} = c_1\boldsymbol{\alpha}_1 + c_2\boldsymbol{\alpha}_2 + \ldots + c_n\boldsymbol{\alpha}_n = \sum_{i=1}^{n} c_i\boldsymbol{\alpha}_i.$$

1.1.1. Ejemplos de un espacio vectorial

Ejemplo 1.1.1 ***El espacio de*** n***-adas***, $\mathbb{R}^n$. *Sea* V *el conjunto de todas las* n*-adas* $\boldsymbol{\alpha} = (x_1, x_2, \ldots, x_n)$ *de escalares* x_i *de* $\mathbb{R}$. *Si* $\boldsymbol{\beta} = (y_1, y_2, \ldots, y_n)$ *con* y_i *en* $\mathbb{R}$, *la suma de* $\boldsymbol{\alpha}$ *y* $\boldsymbol{\beta}$ *se define por*

$$\boldsymbol{\alpha} + \boldsymbol{\beta} = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n).$$

El producto por un escalar $c \in \mathbb{R}$ *y el vector* $\boldsymbol{\alpha}$ *se define por*

$$c\boldsymbol{\alpha} = (cx_1, cx_2, \ldots, cx_n).$$

Ejemplo 1.1.2 ***El espacio de matrices*** $m \times n$, $\mathbb{M}^{m\times n}$. *Sean* m *y* n *enteros positivos. Sea* $\mathbb{M}^{m\times n}$ *el conjunto de todas las matrices* $m \times n$ *sobre* $\mathbb{M}$. *La suma de dos matrices* $\boldsymbol{A}$ *y* $\boldsymbol{B}$ *en* $\mathbb{M}^{m\times n}$ *se define por*

$$(\boldsymbol{A} + \boldsymbol{B})_{ij} = \boldsymbol{A}_{ij} + \boldsymbol{B}_{ij.}$$

El producto por un escalar $c \in \mathbb{R}$ *y la matriz* $\boldsymbol{A}$ *se define por*

$$(c\boldsymbol{A})_{ij} = c\boldsymbol{A}_{ij}.$$

1.2. Subespacio vectorial

En esta sección se introducirá el concepto de subespacio vectorial, así como algunos conceptos relacionados con los espacios vectoriales, además de algunos de sus resultados.

Definición 1.2.1 *Sea* V *un espacio vectorial sobre* $\mathbb{R}$. *Un* ***subespacio*** *de* V *es un subconjunto* W *de* V *que, con las operaciones de adición vectorial y multiplicación por un escalar sobre* V, *es él mismo un espacio vectorial sobre* $\mathbb{R}$.

Teorema 1.2.1 *Un subconjunto no vacío* S *de* V *es un subespacio de* V *si y sólo si para cada par de vectores* $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ *en* S *y cada escalar* c *en* $\mathbb{R}$ *el vector* $c\boldsymbol{\alpha} + \boldsymbol{\beta}$ *está en* S.

Demostración: Ver [6], pag 35.

Teorema 1.2.2 *Sea V un espacio vectorial sobre $\mathbb{R}$. La intersección de cualquier colección de subespacios de V es un subespacio de V.*

Demostración: Ver [6], pag 36.

Definición 1.2.2 *Sea S un conjunto de vectores de un espacio vectorial V. El* ***subespacio generado*** *por S se define como la intersección W de todos los subespacios de V que contienen a S. Cuando S es un conjunto finito de vectores, $S = \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n\}$ se dice simplemente que W es el* ***subespacio generado por los vectores*** *$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$.*

Teorema 1.2.3 *El subespacio generado por un subconjunto S no vacío de un espacio vectorial V es el conjunto de todas las combinaciones lineales de los vectores de S.*

Demostración: Ver [6], pag 37.

1.2.1. Ejemplos de subespacios vectoriales

Ejemplo 1.2.1 *Si V es cualquier espacio vectorial, el subconjunto que consta de solamente el vector nulo es un subespacio vectorial de V, llamado* ***subespacio nulo*** *de V. Es decir*

$$W = \{(x_1, \ldots, x_n) \in \mathbb{R}^n | x_i = 0\}.$$

Ejemplo 1.2.2 *En $\mathbb{R}^n$, el conjunto de las n-adas $(x_1, \ldots, x_n)$ con $x_1 = 0$ es un subespacio, pero el conjunto de las n-adas con $x_1 = 1 + x_2$ no es un subespacio.*

$$W = \{(x_1, \ldots, x_n) \in \mathbb{R}^n | x_1 = 1 + x_2\}.$$

Ejemplo 1.2.3 *El espacio de las funciones polinomios sobre $\mathbb{R}$ es un subespacio del espacio de todas las funciones de $\mathbb{R}$ en $\mathbb{R}$, con las operaciones usuales.*

1.3. Dependencia lineal de vectores

Aquí se definirá la independecia de vectores y la dimensión de un espacio vectorial, mediante la definición de base del mismo.

El concepto de combinación lineal antes mencionado, nos permite dar la definición de dependencia e independencia lineal.

Definición 1.3.1 *Sea V un espacio vectorial sobre $\mathbb{R}$. Un subconjunto S de vectores de V se dice que es* ***linealmente dependiente*** *si existen distintos vectores $\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}, \ldots, \boldsymbol{\alpha_n}$ en S y $c_1, c_2, \ldots, c_n$ escalares en $\mathbb{R}$, no todos iguales a 0, tales que*

$$c_1\boldsymbol{\alpha_1} + c_2\boldsymbol{\alpha_2} + \cdots + c_n\boldsymbol{\alpha_n} = \boldsymbol{0}.$$

Si esto no se cumple, se dice que S es **linealmente independiente**. Es decir, si

$$c_1\boldsymbol{\alpha_1} + c_2\boldsymbol{\alpha_2} + \cdots + c_n\boldsymbol{\alpha_n} = \mathbf{0}$$
$$\Rightarrow c_1 = c_2 = \cdots = c_n = 0.$$

Nótese que cualquier conjunto de vectores que contenga al elemento neutro es linealmente dependiente.

Definición 1.3.2 *Sea V un espacio vectorial. Una* ***base*** *para V es un conjunto de vectores linealmente independiente en V que genera al espacio V.*

Definición 1.3.3 *La* ***dimensión*** *de un espacio vectorial de dimensión finita es el número de elementos de una base cualquiera de V*

El espacio V es de **dimensión finita** si tiene una base finita.

Proposición 1.3.1 *Sea V un espacio vectorial de dimensión n. Entonces:*

1. *Todo conjunto de n vectores linealmente independientes de V es una base de V.*
2. *Todo sistema generador de V de n elementos, es decir, todo conjunto de n vectores pertenecientes a V a partir de los cuales se puede generar V; es una base de V.*

1.4. Producto interno

En esta sección trabajaremos con espacios vectoriales reales. Se dará la definición de producto interno.

Definición 1.4.1 *Sea V un espacio vectorial sobre $\mathbb{R}$. Un* ***producto interno*** *sobre V es una función que asigna a cada par ordenado de vectores $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ de V un escalar $(\boldsymbol{\alpha}|\boldsymbol{\beta})$ en $\mathbb{R}$ de tal modo que para cualesquiera $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ de V y todos los escalares c*

1. $(\boldsymbol{\alpha}+\boldsymbol{\beta}|\boldsymbol{\gamma}) = (\boldsymbol{\alpha}|\boldsymbol{\gamma}) + (\boldsymbol{\beta}|\boldsymbol{\gamma})$;
2. $(c\boldsymbol{\alpha}|\boldsymbol{\beta}) = c(\boldsymbol{\alpha}|\boldsymbol{\beta})$;
3. $(\boldsymbol{\alpha}|\boldsymbol{\alpha}) > 0$ *si* $\boldsymbol{\alpha} \neq 0$.

En $\mathbb{R}^n$ existe un producto interno llamado **producto interno canónico o producto escalar** y se denota por $\langle \boldsymbol{\alpha}, \boldsymbol{\beta} \rangle$, este producto es el que utilizaremos a lo largo del trabajo. Está definido sobre $\boldsymbol{\alpha} = (x_1, \ldots, x_n)$ y $\boldsymbol{\beta} = (y_1, \ldots, y_n)$ por

$$\langle \boldsymbol{\alpha}, \boldsymbol{\beta} \rangle = \sum_j x_i y_i.$$

Definición 1.4.2 *Un* ***espacio producto interno*** *es un espacio real junto con un producto interno definido sobre ese espacio.*

1.5. Matriz no singular y matriz definida positiva

En esta sección se verán algunas definiciones y resultados de las matrices que serán de importancia a lo largo de este trabajo. En general trabajaremos con matrices reales.

1.5.1. Matriz no singular

Definición 1.5.1 *El número de líneas de una matriz $\boldsymbol{A}$ (filas o columnas) que son linealmente independientes, se le conoce como el* ***rango*** *de la matriz $\boldsymbol{A}$ y se denota por $r(\boldsymbol{A})$.*

Definición 1.5.2 *Una matriz $\boldsymbol{A}$ de orden $n \times n$ se dice que es* ***no singular*** *si $r(\boldsymbol{A}) = n$; de otra manera, la matriz es singular.*

Teorema 1.5.1 *Sea* $\boldsymbol{A}$ *una matriz de* $n \times n$. *Entonces las siguientes condiciones son equivalentes:*

1. $\boldsymbol{A}$ *es no singular, i.e.,* $r(\boldsymbol{A}) = n$.
2. *Para cualquier* $\boldsymbol{b} \in \mathbb{R}^n$, $\boldsymbol{Ax} = \boldsymbol{b}$ *tiene un única solución.*
3. *Existe una única matriz* $\boldsymbol{B}$ *tal que* $\boldsymbol{AB} = \boldsymbol{BA} = \boldsymbol{I}$

Demostración: Ver [1], pag 14.

Teorema 1.5.2 *Una matriz es no singular si y sólo si su determinante es diferente de cero.*

Demostración: Ver [1], pag 14.

1.5.2. Matriz definida positiva

Definición 1.5.3 *Una matriz cuadrada* $\boldsymbol{A}$ *es llamada* ***simétrica*** *si* $\boldsymbol{A} = \boldsymbol{A^t}$, *donde* $\boldsymbol{A^t}$ *denota la transpuesta de* $\boldsymbol{A}$.

Definición 1.5.4 *Una matriz* $\boldsymbol{A}$ $n \times n$ *se dice que es definida positiva si es simétrica y si para cualquier vector no nulo* $\boldsymbol{x}$, $\boldsymbol{x^t Ax} > 0$.

Teorema 1.5.3 *Si* $\boldsymbol{A}$ *es definida positiva, entonces es no singular.*

Demostración: Ver [1], pag 18.

Si los menores principales son mayores o iguales que 0, entonces la matriz es semidefinida positiva. Veamos algunas propiedades de estas matrices.

Teorema 1.5.4 *Si* $\boldsymbol{A}$ *es definida positiva entonces existe una matriz no singular* $\boldsymbol{P}$ *tal que* $\boldsymbol{P}^t\boldsymbol{AP} = \boldsymbol{I}$.

Teorema 1.5.5 $\boldsymbol{A}$ *es definida positiva si y sólo si existe* $\boldsymbol{Q}$ *no singular tal que* $\boldsymbol{A} = \boldsymbol{Q}^t\boldsymbol{Q}$.

1.6. Suma directa

Dados dos subespacios $\mathbf{U}$ y $\mathbf{W}$ de un espacio $\mathbf{V}$, cada vector $\boldsymbol{\gamma} \in \mathbf{U} + \mathbf{W}$, tiene descomposición $\boldsymbol{\gamma} = \boldsymbol{\alpha} + \boldsymbol{\beta}$, con $\boldsymbol{\alpha} \in \mathbf{U}$ y $\boldsymbol{\beta} \in \mathbf{W}$, pero en general no es única. En esta sección veremos la importancia de trabajar con sumas directas.

Definición 1.6.1 *Sean $V_1, V_2, V_3, \ldots, V_k$ subespacios del espacio $\mathbb{R}^n$. El **subespacio suma lineal** de $V_1, V_2, V_3, \ldots, V_k$ se define como*

$$V_1 + V_2 + \ldots + V_k = \{\boldsymbol{\alpha} \in \mathbb{R}^n \,|\boldsymbol{\alpha} = \boldsymbol{\alpha_1} + \ldots + \boldsymbol{\alpha_k}, \quad \textit{donde} \quad \boldsymbol{\alpha_i} \in V_i\}.$$

Definición 1.6.2 *Sean $V_1, V_2, V_3, \ldots, V_k$ subespacios del espacio $\mathbb{R}^n$. Se dice que los subespacios $V_1, V_2, V_3, \ldots, V_k$ son **independientes** si*

$$\boldsymbol{\alpha_1} + \ldots + \boldsymbol{\alpha_k} = \boldsymbol{0} \quad \textit{donde} \quad \boldsymbol{\alpha_i} \in V_i,$$

implica que cada $\boldsymbol{\alpha_i}$ es el vector $\boldsymbol{0}$.

Teorema 1.6.1 *Sean $V_1, V_2, V_3, \ldots, V_k$ subespacios del espacio $\mathbb{R}^n$ y sea $V_1 + V_2 + \ldots + V_k$ su suma lineal. Entonces las siguientes afirmaciones son equivalentes:*

1. *Los subespacios $V_1, V_2, \ldots, V_k$ son independientes;*
2. *Se cumple la relación*

$$V_j \cap (V_1 + \ldots + V_{j-1} + V_{j+1} + \ldots + V_K) = \{\boldsymbol{0}\},$$

para todo j tal que $1 \leq j \leq k$.

Demostración: Supóngase que los subespacios $V_1, V_2, \ldots, V_k$ son independientes y sea

$$\boldsymbol{\alpha} \in V_j \cap (V_1 + \ldots + V_{j-1} + V_{j+1} + \ldots + V_k),$$

por lo que,

$$\boldsymbol{\alpha} = \boldsymbol{\alpha_j} \quad \text{y} \quad \boldsymbol{\alpha} = \boldsymbol{\alpha_1} + \ldots + \boldsymbol{\alpha_{j-1}} + \boldsymbol{\alpha_{j+1}} + \ldots + \boldsymbol{\alpha_k}.$$

Entonces tenemos

$$\mathbf{0} = \boldsymbol{\alpha} - \boldsymbol{\alpha} = \boldsymbol{\alpha_1} + \ldots + \boldsymbol{\alpha_{j-1}} + -\boldsymbol{\alpha_j} + \boldsymbol{\alpha_{j+1}} + \ldots + \boldsymbol{\alpha_k}$$

y, como por hipótesis los subespacios son independientes tenemos que cada vector $\boldsymbol{\alpha_i}$ es el vector $\mathbf{0}$, entonces $\boldsymbol{\alpha} = \mathbf{0}$.

Esto implica que

$$V_j \cap (V_1 + \ldots + V_{j-1} + V_{j+1} + \ldots + V_k) = \{\mathbf{0}\}.$$

Ahora, supongase que se cumple la relación

$$V_j \cap (V_1 + \ldots + V_{j-1} + V_{j+1} + \ldots + V_k) = \{\mathbf{0}\}$$

y se tiene

$$\boldsymbol{\alpha_1} + \ldots + \boldsymbol{\alpha_k} = \mathbf{0} \quad \text{con} \quad \boldsymbol{\alpha_i} \in V_i.$$

Sea p el mayor entero j tal que $\boldsymbol{\alpha_j}$ es distinto del vector $\mathbf{0}$. Entonces

$$\boldsymbol{\alpha_p} = (-\boldsymbol{\alpha_1}) + \ldots + (-\boldsymbol{\alpha_{p-1}}) + (-\boldsymbol{\alpha_{p+1}}) + \ldots + (-\boldsymbol{\alpha_k}),$$

que es un vector no nulo en

$$V_p \cap (V_1 + \ldots + V_{p-1} + V_{p+1} + \ldots + V_k) = \{\mathbf{0}\},$$

lo cual nos lleva a una contradicción. Por lo tanto, $\boldsymbol{\alpha_j} = 0$ para todo j $(1 \leq j \leq k)$, y concluimos que los k subespacios $V_1, V_2, \ldots, V_k$ son independientes. ■

Como consecuencia del Teorema 1.5.1 se tiene que si $V_1, V_2, \ldots, V_k$ son subespacios independientes, entonces tales subespacios cumplen la relación $V_i \cap V_j = \{0\}$, si $i \neq j$. La suma lineal de subespacios independientes recibe un nombre especial.

Definición 1.6.3 *Sean $V_1, V_2, \ldots, V_k$ subespacios del espacio $\mathbb{R}^n$. El subespacio suma lineal $V_1 + V_2 + \ldots + V_k$ se denomina **suma directa** de los subespacios $V_1, V_2, \ldots, V_k$, si se cumple cualesquiera de las dos condiciones del Teorema 1.5.1. En este caso denotaremos el subespacio suma lineal por*

$$V_1 \oplus V_2 \oplus \ldots \oplus V_k.$$

Al tener el subespacio suma directa de subespacios, cada uno de los vectores de la suma directa tendrá descomposición única, como lo muestra el siguiente teorema.

Teorema 1.6.2 *Sean $V_1, V_2, \ldots, V_k$ subespacios del espacio $\mathbb{R}^n$. La suma lineal de estos subespacios es suma directa si y sólo si cualquier vector $\boldsymbol{\beta} \in V_1 + V_2 + \ldots + V_k$ tiene descomposición única de la forma*

$$\boldsymbol{\beta} = \boldsymbol{\alpha_1} + \ldots + \boldsymbol{\alpha_k} \quad \textit{donde} \quad \boldsymbol{\alpha_i} \in V_i.$$

Demostración: Supóngase que los subespacios $V_1, V_2, \ldots, V_k$ son independientes y que el vector $\boldsymbol{\beta}$ tiene además la descomposición

$$\boldsymbol{\beta} = \boldsymbol{\alpha'_1} + \ldots + \boldsymbol{\alpha'_k} \quad \text{donde} \quad \boldsymbol{\alpha'_i} \in V_i,$$

formemos la diferencia

$$\mathbf{0} = \boldsymbol{\beta} - \boldsymbol{\beta} = (\boldsymbol{\alpha_1} - \boldsymbol{\alpha'_1}) + \ldots + (\boldsymbol{\alpha_k} - \boldsymbol{\alpha'_k}) \quad \text{donde} \quad (\boldsymbol{\alpha_i} - \boldsymbol{\alpha'_i}) \in V_i.$$

Como los subespacios $V_1, V_2, \ldots, V_k$ son independientes se tiene que

$$(\boldsymbol{\alpha_1} - \boldsymbol{\alpha'_1}) = \ldots = (\boldsymbol{\alpha_k} - \boldsymbol{\alpha'_k}) = \mathbf{0},$$

entonces

$$\boldsymbol{\alpha_i} = \boldsymbol{\alpha'_i}, \quad \forall i \qquad (i = 1, \ldots, k),$$

por lo que la descomposición del vector $\boldsymbol{\beta}$ es única.

Ahora, supóngase que cualquier vector $\boldsymbol{\beta}$ de $V_1+V_2+\ldots+V_k$ tiene una única descomposición de la forma

$$\boldsymbol{\beta} = \boldsymbol{\alpha_1} + \ldots + \boldsymbol{\alpha_k} \quad \text{donde} \quad \boldsymbol{\alpha_i} \in V_i.$$

Sea $\boldsymbol{\alpha_j} \in V_j$ para $j = 1, \ldots, k$, y sea $\boldsymbol{\alpha_1} + \ldots + \boldsymbol{\alpha_k} = \mathbf{0}$. Puesto que $\mathbf{0} \in V_j$ para cada j y $\mathbf{0} = \mathbf{0}+\mathbf{0}+\ldots+\mathbf{0}$, se sique que $\boldsymbol{\alpha_j} = \mathbf{0}$ para todo j, implicando la independencia de los subespacios $V_1, V_2, \ldots, V_k$. ■

Definición 1.6.4 *Sean V_1 y V_2 subespacios de $\mathbb{R}^n$, decimos que V_1 y V_2 son **ortogonales** si para cualquier par de vectores $\boldsymbol{\alpha}$ y $\boldsymbol{\beta}$ tal que $\boldsymbol{\alpha} \in V_1$ y $\beta \in V_2$ se tiene $\langle \boldsymbol{\alpha}, \boldsymbol{\beta} \rangle = 0$. Denotaremos que V_1 y V_2 son ortogonales por $V_1 \perp V_2$.*

En particular, dado cualquier subespacio V de $\mathbb{R}^n$, podemos determinar todos los vectores ortogonales a V. Este conjunto de vectores recibe el siguiente nombre.

Definición 1.6.5 *Sea V un subespacio de $\mathbb{R}^n$. Se define el subespacio* ***complemento ortogonal*** *del subespacio V, como el subespacio de todos los vectores $\boldsymbol{\beta} \in \mathbb{R}^n$, tales que $\langle \boldsymbol{\alpha}, \boldsymbol{\beta} \rangle = 0$, para todo $\boldsymbol{\alpha} \in V$. Denotamos el complemento ortogonal de V, por $V^{\perp}$.*

Si $V_1, V_2, \ldots, V_k$ son subespacios de $\mathbb{R}^n$, tales que $V_i \perp V_j$ si $i \neq j$, decimos que son mutumente ortogonales.

Teorema 1.6.3 *Sean $V_1, V_2, \ldots, V_k$ subespacios del espacio $\mathbb{R}^n$ mutuamente ortogonales, entonces su suma lineal es directa.*

Demostración: Supóngase

$$\boldsymbol{\alpha} \in V_j \cap (V_1 + \ldots + V_{j-1} + V_{j+1} + \ldots + V_k).$$

Así, $\boldsymbol{\alpha} \in V_j$ y $\boldsymbol{\alpha} \in (V_1 + \ldots + V_{j-1} + V_{j+1} + \ldots + V_k)$
lo cual implica

$$\langle \boldsymbol{\alpha}, \boldsymbol{\alpha} \rangle = \langle \boldsymbol{\alpha_j}, \boldsymbol{\alpha_1} \rangle + \ldots + \langle \boldsymbol{\alpha_j}, \boldsymbol{\alpha_{j-1}} \rangle + \langle \boldsymbol{\alpha_j}, \boldsymbol{\alpha_{j+1}} \rangle + \ldots + \langle \boldsymbol{\alpha_j}, \boldsymbol{\alpha_k} \rangle,$$

pero por hipótesis los subespacios son ortogonales, es decir, $\langle \boldsymbol{\alpha_j}, \boldsymbol{\alpha_p} \rangle = 0$, para todo $p = 1, \ldots, k$ con $p \neq j$, así tenemos que $\langle \boldsymbol{\alpha}, \boldsymbol{\alpha} \rangle = 0$, de lo que se sigue que $\boldsymbol{\alpha} = \mathbf{0}$, y así

$$V_j \cap (V_1 + \ldots + V_{j-1} + V_{j+1} + \ldots + V_k) = \{\mathbf{0}\}.$$

Por lo tanto, finalmente tenemos el subespacio suma directa de los subespacios $V_1, V_2, \ldots, V_k$. $\blacksquare$

Capítulo 2

OPERADOR PROYECTOR Y SUS PROPIEDADES

En este capítulo introduciremos el concepto de transformación lineal, así como los de operador lineal y operador proyector, incluyendo algunas de sus propiedades.

2.1. Transformaciones lineales

Aquí veremos la definición de una transformación lineal, la cuál nos permitirá definir un operador lineal. Sean V y W espacios vectoriales sobre $\mathbb{R}$. Una

transformación lineal de V en W es una función T de V en W tal que

$$T(c\boldsymbol{\alpha} + \boldsymbol{\beta}) = c(T(\boldsymbol{\alpha})) + T(\boldsymbol{\beta})$$

para todo $\boldsymbol{\alpha}$ y $\boldsymbol{\beta}$ en V y para todo escalar c en $\mathbb{R}$.

Proposición 2.1.1 *Sea $T : V \to W$ una transformación lineal. Entonces:*

1. *Si V_1 es un subespacio de V, entonces $T(V_1)$ es un subespacio de W.*
2. *Si W_1 es un subespacio de W, entonces $T^{-1}(W_1)$ es un subespacio de V.*

Teorema 2.1.1 *Sean V y W espacios vectoriales sobre $\mathbb{R}$. Sean T y U transformaciones lineales de V en W. La función $(T+U)(\boldsymbol{\alpha})$ definida por*

$$(T+U)(\boldsymbol{\alpha}) = T(\boldsymbol{\alpha}) + U(\boldsymbol{\alpha})$$

es una transformación lineal de V en W. Si c es cualquier elemento de $\mathbb{R}$ la función (cT) definida por

$$(cT)(\alpha) = c(T(\alpha))$$

es una transformación lineal de V en W, el conjunto de todas las transformaciones lineales de V en W, junto con la adición y la multiplicación por un escalar aquí definidas, es un espacio vectorial sobre $\mathbb{R}$.

Demostración: Ver [6], pag 74.

Definición 2.1.1 *Si V es un espacio vectorial en $\mathbb{R}$, un **operador lineal** sobre V es una transformación lineal de V en V.*

2.1.1. Ejemplos de transformaciones lineales

Ejemplo 2.1.1 *Si V es cualquier espacio vectorial, la **transformación identidad $\boldsymbol{I}$**, definida por $\boldsymbol{I\alpha} = \boldsymbol{\alpha}$, es una transformación lineal de V en V.*

Ejemplo 2.1.2 *Sea $\boldsymbol{A}$ una matriz $m \times n$ dada, con elementos en $\mathbb{R}$. La función T definida por $T(\boldsymbol{X}) = \boldsymbol{AX}$ es una transformación lineal de $\mathbb{R}^{n\times 1}$ en $\mathbb{R}^{m\times 1}$. La función U definida por $U(\boldsymbol{\alpha}) = \boldsymbol{\alpha}A$ es una transformación lineal de $\mathbb{R}^m$ en $\mathbb{R}^n$.*

2.1.2. Traza y rango de una matriz

Aquí se introduciran los conceptos de traza y rango de una matriz, los cuales utilizaremos a lo largo de este trabajo.

Definición 2.1.2 *Si $\boldsymbol{A}$ es una matriz $n \times n$ con elementos en $\mathbb{R}$, la **traza** de $\boldsymbol{A}$ es el escalar*

$$tr(\boldsymbol{A}) = \sum_{i=1}^{n} a_{ii}$$

donde los a_{ii} son los elementos de la matriz que están sobre la diagonal.

2.2. Matriz asociada a una transformación lineal

En esta sección veremos como a cada transformación lineal en algún espacio V, se le asocia un matriz respecto a sus bases.

Si V y W son espacios vectoriales de dimensión n y m respectivamente, una transformación lineal $T : V \to W$ queda unívocamene determinada por los n vectores de W que son los valores de T en una base de V. Se define una matriz asociada a T que contiene toda esta información. T está determinada por su efecto sobre los vectores $\boldsymbol{\alpha_j}$. Cada uno de los n vectores $T(\boldsymbol{\alpha_j})$ se expresan de manera única como combinación lineal

$$T(\boldsymbol{\alpha_j}) = \sum_{i=1}^{m} A_{ij}\beta_i,$$

de los β_i, los escalares $A_{1j}, \ldots, A_{mj}$ son las coordenadas de $T(\boldsymbol{\alpha_j})$ en la base ordenada $\boldsymbol{\beta_W}$.

Teorema 2.2.1 *Sea V un espacio vectorial de dimensión n sobre $\mathbb{R}$, y W un espacio vectorial de dimensión m sobre $\mathbb{R}$. Sea $\boldsymbol{\beta_V}$ una base ordenada de V y $\boldsymbol{\beta_W}$ una base ordenada de W. Para cada transformación lineal T de V en W, existe una matriz $m \times n$, A, cuyos elementos pertenecen a $\mathbb{R}$, tal que*

$$[T(\boldsymbol{\alpha})]_{\boldsymbol{\beta_W}} = A[\boldsymbol{\alpha}]_{\boldsymbol{\beta_V}}$$

para todo vector $\boldsymbol{\alpha}$ en V. Además, $T \to A$ es una correspondecia biyectiva entre el conjunto de todas las transformaciones lineales de V en W y el conjunto de todas las matrices de $m \times n$ sobre $\mathbb{R}$.

Demostración: Ver [6], pag 86.

Definición 2.2.1 *La matriz A, que está asociada a T en el teorema anterior se llama la* ***matriz de T respecto a las bases $\boldsymbol{\beta_V}$*** *y* $\boldsymbol{\beta_W}$.

Obsérvese que A es la matriz cuyas columnas $A_1, \ldots, A_n$ son dadas por

$$A_j = [T\boldsymbol{\alpha_j}]_{\boldsymbol{\beta_W}}.$$

2.3. Operador proyector

Aquí se trabajará con algunas propiedades del operador proyector.

Sean V_1 y V_2 subespacios independientes de $\mathbb{R}^n$, tales que $V_1 \oplus V_2 = \mathbb{R}^n$. Por la sección de suma directa, tenemos que cualquier elemento $\boldsymbol{\alpha} \in \mathbb{R}^n$, tiene descomposición única $\boldsymbol{\alpha} = \boldsymbol{\alpha_1} + \boldsymbol{\alpha_2}$, donde $\boldsymbol{\alpha_1} \in V_1$ y $\boldsymbol{\alpha_2} \in V_2$. Denotemos por P al mapeo de $\mathbb{R}^n$ a V_1 dado por $P(\boldsymbol{\alpha}) = \boldsymbol{\alpha_1}$. El operador P es un operador lineal sobre $\mathbb{R}^n$.

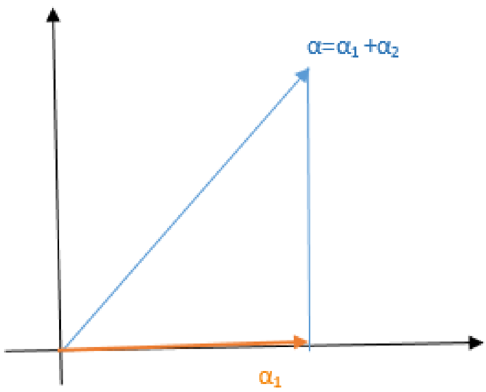

Figura 2.1: Operador proyector $P(\boldsymbol{\alpha}) = \boldsymbol{\alpha_1}$.

Definición 2.3.1 *Al operador P se le denomina el* ***operador proyector*** *sobre V_1 a lo largo de V_2.*

A continuanción presentaremos algunas propiedades del operador proyector.

Teorema 2.3.1 *Un operador P sobre $\mathbb{R}^n$ es un operador proyector sobre algún subespacio V_1 si y sólo si, es idempotente, es decir, si $PP = P$.*

Demostración: Sea P el operador proyector sobre V_1 a lo largo de V_2. Si $\boldsymbol{\alpha_1} \in V_1$, la proyección de $\boldsymbol{\alpha_1}$ sobre V_1 a lo largo de V_2 es $\boldsymbol{\alpha_1}$. Para cualquier $\boldsymbol{\alpha} \in \mathbb{R}^n$, se tiene $\boldsymbol{\alpha} = \boldsymbol{\alpha_1} + \boldsymbol{\alpha_2}$, donde $\boldsymbol{\alpha_1} \in V_1$ y $\boldsymbol{\alpha_2} \in V_2$. Tenemos que $PP(\boldsymbol{\alpha}) = P(\boldsymbol{\alpha_1}) = \boldsymbol{\alpha_1} = P(\boldsymbol{\alpha})$, por lo que, P es un operador idempotente.

Inversamente, sea P un operador idempotente. Definamos V_1 como el conjunto de todos los vectores $\boldsymbol{\alpha} \in \mathbb{R}^n$, tales que $\boldsymbol{\alpha} = P(\boldsymbol{\alpha})$ y sea V_2 el conjunto de todos los vectores $\boldsymbol{\alpha} \in \mathbb{R}^n$ tales que $P(\boldsymbol{\alpha}) = 0$; es decir,

$$\begin{aligned} V_1 &= \{\boldsymbol{\alpha} \in \mathbb{R}^n | P(\boldsymbol{\alpha}) = \boldsymbol{\alpha}\}, \\ V_2 &= \{\boldsymbol{\alpha} \in \mathbb{R}^n | P(\boldsymbol{\alpha}) = 0\}. \end{aligned}$$

Ahora, supóngase que $\boldsymbol{\alpha} \in V_1 \cap V_2$, entonces $\boldsymbol{\alpha} = 0$, así, $V_1 \cap V_2 = \{0\}$; además, para cualquier $\boldsymbol{\alpha} \in \mathbb{R}^n$, se tiene que $\boldsymbol{\alpha} = P(\boldsymbol{\alpha}) + (I - P)\boldsymbol{\alpha}$. Definamos $\boldsymbol{\alpha_1} = P(\boldsymbol{\alpha})$ y $\boldsymbol{\alpha_2} = (I - P)\boldsymbol{\alpha}$, entonces $\boldsymbol{\alpha} = \boldsymbol{\alpha_1} + \boldsymbol{\alpha_2}$ y se tiene que

$$P(\boldsymbol{\alpha_1}) = PP(\boldsymbol{\alpha}) = P(\boldsymbol{\alpha}) = \boldsymbol{\alpha_1}$$

por lo que $\boldsymbol{\alpha_1} \in V_1$, y

$$P(\boldsymbol{\alpha_2}) = 0,$$

lo cual implica que $\boldsymbol{\alpha_2} \in V_2$. Así, $V_1 + V_2 = \mathbb{R}^n$. Así tenemos que $V_1 \oplus V_2 = \mathbb{R}^n$. Además, como $\boldsymbol{\alpha} = \boldsymbol{\alpha_1} + \boldsymbol{\alpha_2}$, se tiene que

$$P(\boldsymbol{\alpha}) = P(\boldsymbol{\alpha_1}) + P(\boldsymbol{\alpha_2}) = \boldsymbol{\alpha_1} + 0 = \boldsymbol{\alpha_1}. \tag{2.1}$$

Entonces, por la definición de operador proyector se tiene de (2.1) que P es el operador proyector sobre V_1 a lo largo de V_2. ■

Teorema 2.3.2 *Sean V_1 y V_2 dos subespacios independientes de $\mathbb{R}^n$, tales que $V_1 \oplus V_2 = \mathbb{R}^n$, y sea P un operador lineal sobre $\mathbb{R}^n$. Tenemos que P es el operador proyector sobre V_1 a lo largo de V_2 si y sólo si P satisface*

$$P(\boldsymbol{\alpha}) = \begin{cases} \boldsymbol{\alpha} & \text{si } \boldsymbol{\alpha} \in V_1 \\ 0 & \text{si } \boldsymbol{\alpha} \in V_2 \end{cases} \tag{2.2}$$

Demostración: Denotemos con $P_{1,2}$ el operador proyector sobre V_1 a lo largo de V_2 y con $P_{2,1}$ el operador proyector sobre V_2 a lo largo de V_1. Sea cualquier $\boldsymbol{\alpha} \in \mathbb{R}^n$, con $\boldsymbol{\alpha} = \boldsymbol{\alpha_1} + \boldsymbol{\alpha_2}$, en donde $\boldsymbol{\alpha_1} \in V_1$ y $\boldsymbol{\alpha_2} \in V_2$. Ahora, $\boldsymbol{\alpha} = \boldsymbol{\alpha_1} + \boldsymbol{\alpha_2} = P_{1,2}(\boldsymbol{\alpha_1}) + P_{2,1}(\boldsymbol{\alpha_2})$, de lo cual obtenemos $P(\boldsymbol{\alpha}) = P(P_{1,2}(\boldsymbol{\alpha_1})) + P(P_{2,1}(\boldsymbol{\alpha_2}))$. Además, tenemos que $P_{1,2}(\boldsymbol{\alpha_1}) \in V_1$ y $P_{2,1}(\boldsymbol{\alpha_2}) \in V_2$; así, por la condición (2.2) tenemos que

$$P(\boldsymbol{\alpha}) = P_{1,2}(\boldsymbol{\alpha_1}) = \boldsymbol{\alpha_1};$$

así $P(\boldsymbol{\alpha}) = \boldsymbol{\alpha_1}$, con lo que se tiene, por definición de operdor preyector, que P es el operador proyector sobre V_1 a lo largo de V_2.

Inversamente, sea P el operador proyector sobre V_1 a lo largo de V_2, y sean los vectores $\boldsymbol{\alpha_1} \in V_1$ y $\boldsymbol{\alpha_2} \in V_2$, así se tiene que $P(\boldsymbol{\alpha_1}) = P(\boldsymbol{\alpha_1} + 0) = \boldsymbol{\alpha_1}$ y $P(\boldsymbol{\alpha_2}) = P(0 + \boldsymbol{\alpha_2}) = 0$. Entonces, P satisface

$$P(\boldsymbol{\alpha}) = \begin{cases} \boldsymbol{\alpha} & \text{si } \boldsymbol{\alpha} \in V_1 \\ 0 & \text{si } \boldsymbol{\alpha} \in V_2. \end{cases}$$

Notemos que si P es el operador proyector sobre V_1 a lo largo de V_2, entonces por el Teorema 2.3.2 tenemos que $(I - P)$ es el operador proyector sobre V_2 a lo largo de V_1. Además, $P(I - P) = (I - P)P = 0$.

Ahora veremos la relación entre proyectores y subespacios.

Teorema 2.3.3 *Sean V_1 y V_2 subespacios del espacio $\mathbb{R}^n$, tales que $V_1 \subset V_2$, con subespacios complementarios W_1 y W_2, respectivamente; es decir, $V_1 \oplus W_1 = \mathbb{R}^n$ y $V_2 \oplus W_2 = \mathbb{R}^n$. Sean P_1 y P_2 los proyectores sobre V_1 a lo largo de W_1 y sobre V_2 a lo largo de W_2, respectivamente, entonces*

$$P_1P_2(\boldsymbol{\alpha}) = P_2P_1(\boldsymbol{\alpha}) = P_1(\boldsymbol{\alpha}).$$

Demostración: Sean $Q_1 = I - P_1$, el proyector sobre W_1 a lo largo de V_1, y $Q_2 = I - P_2$ el proyector sobre W_2 a lo largo de V_2. Sea cualquier vector $\boldsymbol{\alpha} \in \mathbb{R}^n$, como por hipótesis $V_1 \subset V_2$, tenemos que $P_1(\boldsymbol{\alpha}) \in V_2$, entonces

$$P_2P_1 = P_1. \tag{2.3}$$

Ahora, por hipótesis también se cumple que $W_2 \subset W_1$, entonces de (2.3) tenemos que $Q_1Q_2 = Q_2$, lo cual implica

$$(I - P_1)(I - P_2) = I - P_2,$$

obteniéndose

$$-P_1 + P_1P_2 = 0,$$

de lo cual tenemos que

$$P_1 = P_1P_2. \tag{2.4}$$

Así, de (2.3) y de (2.4), obtenemos

$$P_1P_2 = (P_2P_1)P_2 = P_2(P_1P_2) = P_2P_1.$$

Así, tenemos que $P_1P_2 = P_2P_1$. ■

Ahora presentaremos una clase especial del operador proyector.

Definición 2.3.2 *Sea V un subespacio del espacio $\mathbb{R}^n$, y sea $V^{\perp}$ su complemento ortogonal. Entonces el operador proyector sobre V a lo largo de $V^{\perp}$, se denomina* ***proyector ortogonal sobre*** *V.*

Se denotará por $S(\mathbf{X})$ al espacio generado por los vectores columna de la matriz $\mathbf{X}$.

Teorema 2.3.4 $\mathbf{P}$ *es el operador proyector ortogonal sobre $S(\mathbf{X})$ sí y sólo sí*

a) $\mathbf{P}(\boldsymbol{\alpha})=\boldsymbol{\alpha}$ *para cualquier* $\boldsymbol{\alpha} \in S(\mathbf{X})$ *y*

b) $\mathbf{P}(\boldsymbol{\beta})=\mathbf{0}$ *para cualquier* $\boldsymbol{\beta} \in S(\mathbf{X})^{\perp}$

Demostración: Se tiene que $S(\mathbf{X}) \oplus S(\mathbf{X})^{\perp} = \mathbb{R}^n$, así por la definición de proyector ortogonal y por el Teorema 2.3.2 se cumple el resultado.

2.4. Matriz inversa generalizada

Definición 2.4.1 *Una matriz inversa generalizada de* $\mathbf{A}$ *es una matriz* $\mathbf{G}$ *que satisface que* $\mathbf{AGA} = \mathbf{A}$.

La notación usada para denotar una inversa generalizada de la matriz $\mathbf{A}$ es $\mathbf{A}^{-}$.

Teorema 2.4.1 *Si* $\mathbf{G}$ *es una matriz inversa generalizada de* $\mathbf{X}^t\mathbf{X}$*, entonces* $\mathbf{G}\mathbf{X}^t$ *es una matriz inversa generalizada de* $\mathbf{X}$*, es decir, se cumple que*

$$\mathbf{X}\mathbf{G}\mathbf{X}^t\mathbf{X} = \mathbf{X}.$$

Demostración: Sea $\alpha \in \mathbb{R}^n$, éste tiene descomposición única $\alpha = \alpha_1 + \alpha_2$, en donde $\alpha_1 \in S(\mathbf{X})$ y $\alpha_2 \in S(\mathbf{X})^\perp$. Así, $\alpha_1 = \mathbf{X}\gamma_1$ para algún γ_1. Entonces

$$\begin{aligned} \alpha^t\mathbf{X}\mathbf{G}\mathbf{X}^t\mathbf{X} &= \alpha_1^t\mathbf{X}\mathbf{G}\mathbf{X}^t\mathbf{X} \\ &= \gamma_1^t\left(\mathbf{X}^t\mathbf{X}\right)\mathbf{G}\left(\mathbf{X}^t\mathbf{X}\right) \\ &= \gamma_1^t\left(\mathbf{X}^t\mathbf{X}\right) \\ &= \alpha^t\mathbf{X}. \end{aligned}$$

Así se cumple, $\mathbf{X}\mathbf{G}\mathbf{X}^t\mathbf{X} = \mathbf{X}$.

Teorema 2.4.2 *Sea* $(\mathbf{X}^t\mathbf{X})^-$ *una inversa generalizada de la matriz* $(\mathbf{X}^t\mathbf{X})$*, entonces* $\mathbf{X}(\mathbf{X}^t\mathbf{X})^-\mathbf{X}^t$ *es el operador proyector ortogonal sobre* $S(\mathbf{X})$*.*

Demostración: Si $\alpha \in S(\mathbf{X})$ entonces $\alpha = \mathbf{X}\gamma$ y por el Teorema 2.4.1 se tiene

$$\begin{aligned} \mathbf{X}\left(\mathbf{X}^t\mathbf{X}\right)^-\mathbf{X}^t(\alpha) &= \mathbf{X}\left(\mathbf{X}^t\mathbf{X}\right)^-\mathbf{X}^t\mathbf{X}\gamma \\ &= \mathbf{X}\gamma = \alpha. \end{aligned}$$

Ahora, si $\beta \in S(\mathbf{X})^\perp$ se tiene que $\mathbf{X}(\mathbf{X}^t\mathbf{X})^-\mathbf{X}^t(\beta)$=$\mathbf{0}$. Así se cumple que

a) $\mathbf{X}(\mathbf{X}^t\mathbf{X})^-\mathbf{X}^t(\alpha)$=$\alpha$ para cualquier $\alpha \in S(\mathbf{X})$ y

b) $\mathbf{X}(\mathbf{X}^t\mathbf{X})^-\mathbf{X}^t(\beta)$=$\mathbf{0}$ para cualquier $\beta \in S(\mathbf{X})^\perp$

y por el Teorema 2.3.4 se tiene el resultado.

Sea $\mathbf{V}$ una matriz definida positiva. Así existe una matriz no singular $\mathbf{R}$ para la cual se cumple $\mathbf{V} = \mathbf{R}\mathbf{R}^t$. Es de interés el operador proyector ortogonal $\mathbf{P}_\mathbf{R}$ sobre $S(\mathbf{R}^{-1}\mathbf{X})$, el cual, por el Teorema 2.4.2 está dado por medio de

$$\left(\mathbf{R}^{-1}\mathbf{X}\right)\left[(\mathbf{R}^{-1}\mathbf{X})^t(\mathbf{R}^{-1}\mathbf{X})\right]^-\left(\mathbf{R}^{-1}\mathbf{X}\right)^t.$$

Teorema 2.4.3 *Sea* $\mathbf{V}$ *una matriz definida positiva, entonces*

$$\mathbf{X}\left(\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}^t\mathbf{V}^{-1}$$

es un operador proyector sobre $S(\mathbf{X})$*.*

Demostración: Denotaremos por $\mathbf{P_{XV}}$ al operador $\mathbf{X}\left(\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}^t\mathbf{V}^{-1}$. En primer lugar veamos que efectivamente es un operador proyector. Se tiene que

$$\begin{aligned}
\mathbf{P_{XV}}^2 &= \left[\mathbf{X}\left(\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}^t\mathbf{V}^{-1}\right]\left[\mathbf{X}\left(\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}^t\mathbf{V}^{-1}\right] \\
&= \left[\mathbf{X}\left(\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}\right)^{-}\left(\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}\right)\left(\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}^t\mathbf{V}^{-1}\right] \\
&= \mathbf{X}\left(\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}^t\mathbf{V}^{-1} \\
&= \mathbf{P_{XV}},
\end{aligned}$$

y por lo tanto se tiene que se cumple la idempotencia, así por el Teorema 2.3.1 $\mathbf{P_{XV}}$ es un operador proyector. Ahora veamos que es un proyector sobre $S(\mathbf{X})$. Se tiene que el operador proyector ortogonal $\mathbf{P_R}$ sobre $S(\mathbf{R}^{-1}\mathbf{X})$, cumple la relación

$$\left(\mathbf{R}^{-1}\mathbf{X}\right)\left[(\mathbf{R}^{-1}\mathbf{X})^t(\mathbf{R}^{-1}\mathbf{X})\right]^{-}\left(\mathbf{R}^{-1}\mathbf{X}\right)^t\left(\mathbf{R}^{-1}\mathbf{X}\right) = \left(\mathbf{R}^{-1}\mathbf{X}\right),$$

lo cual puede ser escrito como:

$$\mathbf{R}^{-1}\mathbf{X}\left(\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X} = \mathbf{R}^{-1}\mathbf{X},$$

de lo que se obtiene que

$$\mathbf{X}\left(\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X} = \mathbf{X}. \tag{2.5}$$

Si $\alpha \in \mathrm{S}(\mathbf{X})$ entonces $\alpha = \mathbf{X}\gamma$. Entonces

$$\begin{aligned}
\mathbf{P_{XV}}(\alpha) &= \mathbf{X}\left(\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}^t\mathbf{V}^{-1}\alpha \\
&= \mathbf{X}\left(\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}\gamma \\
&= \mathbf{X}\gamma \\
&= \alpha.
\end{aligned}$$

Así se cumple que $\mathbf{P_{XV}}(\alpha) = \alpha$ para $\alpha \in S(\mathbf{X})$.

Capítulo 3

OPERADOR PROYECTOR EN EL MODELO LINEAL GENERAL

3.1. Modelo lineal general

Esta sección trata sobre el modelo lineal general, comenzando por la idea de la posible relación lineal entre dos variables, una variable respuesta y la otra una variable explicatoria, esto nos permitirá introducir el modelo lineal general.

El modelo lineal general que es usado para determinar una observación y de una variable Y de algún valor x conocido de una variable X es usualmente escrito de la forma

$$y = f(x) + e$$

en donde e es una variable aleatoria, $f(x)$ es una función de x, y x es una variable que no es aleatoria. La función $f(x)$ es definida como la parte determinista del modelo, Y y e la parte aleatoria. También, Y se refiere como variable dependiente o la variable respuesta, y x es definida como variable independiente o variable predictora. La variable e no es observable, pero algunas cosas acerca de la distribución de e se afirman como parte del modelo.

En el caso más simple $f(x) = \beta_0 + \beta_1 x_1$, β_0 y β_1 son parámetros desconocidos. A menudo β_0 y β_1 pueden ser cualquier número real. Entonces podemos escribir este modelo como

$$Y = \beta_0 + \beta_1 x_1 + e.$$

En su forma más general, $f(x)$ es una función lineal de $k+1$ parámetros desconocidos $\beta_0, \beta_1, \ldots, \beta_k$ y se puede escribir como

$$f(x) = \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k,$$

donde las x_i contienen parámetros desconocidos.

Para la i-ésima observación, el modelo queda dado por:

$$Y_i = \beta_0 + \beta_i x_{1i} + e_i, \qquad E(e_i) = 0, \qquad \text{para } i = 1, 2, \ldots, n.$$

Este conjunto de relaciones es llamado el modelo simple. Con las observaciones podemos calcular las estimaciones de β_0 y β_1.

Definición 3.1.1 *Consideremos las n ecuaciones*

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_k x_{ki} + e_i, \qquad E[e_i] = 0, \qquad i = 1, 2, \ldots, n.$$

en donde

1. *Las Y_i son variables aleatorias observables;*
2. *Las x_{ji} son observables;*
3. *β_0 y β_i para $i = 1, \ldots, k$ son parámetros desconocidos;*
4. *Las e_i son variables aleatorias no observables tales que $cov(e_i, e_i) = \sigma_{ii}$ y $cov(e_i, e_{i'}) = 0$.*

Estas especificaciones definen un modelo lineal simple.

Podemos escribir las n ecuaciones del modelo de la definición anterior como

$$\begin{aligned}
Y_1 &= \beta_0 + \beta_1 x_{11} + \beta_2 x_{21} + \ldots + \beta_k x_{k1} + e_1 \\
Y_2 &= \beta_0 + \beta_1 x_{12} + \beta_2 x_{22} + \ldots + \beta_k x_{k2} + e_2 \\
&\vdots \qquad\qquad\qquad\qquad\qquad\qquad \vdots \\
Y_n &= \beta_0 + \beta_1 x_{1n} + \beta_2 x_{2n} + \ldots + \beta_k x_{kn} + e_n
\end{aligned}$$

o en forma matricial

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & \dots & x_{k1} \\ 1 & x_{12} & \dots & x_{k2} \\ \vdots & \vdots & \dots & \vdots \\ 1 & x_{1n} & \dots & x_{kn} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

Es decir:

$$\mathbf{Y} = X\beta + \mathbf{e}, \qquad E(\mathbf{e}) = \mathbf{0}, \qquad \text{y} \qquad cov(\mathbf{e}) = \sigma^2 \mathbf{I}.$$

3.2. Estimación en el modelo lineal general

Con el fin de estimar β necesitamos determinar un estimador $\overset{\wedge}{\beta}$, el cual es una función de $\mathbf{Y}$ y de otras cantidades como $\mathbf{X}$, tal que $\overset{\wedge}{\beta}$ sea cercano a β en algún sentido. En este caso, si sustituimos $\overset{\wedge}{\beta}$ por β, entonces $\mathbf{Y}$ estará cercano a $\mathbf{X}\overset{\wedge}{\beta}$. La diferencia

$$\mathbf{Y} - \mathbf{X}\overset{\wedge}{\beta} = \epsilon$$

se denomina el vector de residuales, mientras que

$$\mathbf{Y} - \mathbf{X}\beta = \mathbf{e},$$

se denomina el vector de errores. Un método de elegir $\overset{\wedge}{\beta}$ es minimizar la suma de cuadrados de los elementos de $\mathbf{e}$.

Se tiene que $E(\mathbf{Y}) = \mathbf{X}\beta$, pero β es desconocido, lo que se sabe es que $E(\mathbf{Y}) \in S(\mathbf{X})$. Para la estimación de $E(\mathbf{Y})$ se toma el vector perteneciente a $S(\mathbf{X})$ que esté más cercano a $\mathbf{Y}$. La distancia está dada en términos de $(\mathbf{Y} - \mathbf{X}\beta)^t (\mathbf{Y} - \mathbf{X}\beta)$. Así, se tiene la siguiente definición.

Definición 3.2.1 *Un estimador $\overset{\wedge}{\beta}$ se denomina un estimador por mínimos cuadrados de β, si $\mathbf{X}\overset{\wedge}{\beta}$ es el vector perteneciente a $S(\mathbf{X})$ que esta más cercano a $\mathbf{Y}$. En otras palabras, $\overset{\wedge}{\beta}$ es un estimador de mínimos cuadrados si*

$$\left(\mathbf{Y} - \mathbf{X}\overset{\wedge}{\beta}\right)^t \left(\mathbf{Y} - \mathbf{X}\overset{\wedge}{\beta}\right) = \min_{\beta} (\mathbf{Y} - \mathbf{X}\beta)^t (\mathbf{Y} - \mathbf{X}\beta). \tag{3.1}$$

3.3. Operador proyector en la estimación en el modelo lineal general

En esta sección se muestra como el operador proyector ortogonal sobre S($\mathbf{X}$) interviene en la estimación de los parámetros en el modelo lineal general; es decir, en la estimación de los coeficientes de regresión y en la estimacón de la varianza, se presentan los estimadores por mínimos cuadrados y los estimadores por mínimos cuadrados generalizados.

3.3.1. Mínimos cuadrados ordinarios

Uno de los métodos más utilizados para obtener estimaciones de los coeficientes de regresión es el método de mínimos cuadrado, el cual se muestra a continuación.

Considérese el modelo

$$\mathbf{Y} = \mathbf{X}\beta + \mathbf{e}, \qquad E(\mathbf{e}) = \mathbf{0}, \quad Cov(\mathbf{e}) = \sigma^2\mathbf{I}, \tag{3.2}$$

donde $\mathbf{Y} \in \mathbb{R}^n$, $\mathbf{X}$ es una matriz de constantes de orden $n \times p$, $\beta \in \mathbb{R}^p$ es un vector de parámetros desconocidos, y $\mathbf{e} \in \mathbb{R}^n$ es un vector de errores aleatorios no observables.

Teorema 3.3.1 *Sea $\mathbf{P_X}$ el operador proyector ortogonal sobre $S(\mathbf{X})$, entonces*

$$(\mathbf{Y} - \mathbf{X}\beta)^t (\mathbf{Y} - \mathbf{X}\beta) = (\mathbf{Y} - \mathbf{P_X Y})^t (\mathbf{Y} - \mathbf{P_X Y}) + (\mathbf{P_X Y} - \mathbf{X}\beta)^t (\mathbf{P_X Y} - \mathbf{X}\beta).$$

Demostración: Lo anterior se demuestra en base a las propiedades del operador proyector ortogonal, $(\mathbf{I} - \mathbf{P_X})\,\mathbf{P_X} = \mathbf{0}$ y $(\mathbf{I} - \mathbf{P_X})\,\mathbf{X} = \mathbf{0}$. ■

Teorema 3.3.2 *$\overset{\wedge}{\beta}$ es un estimador por mínimos cuadrados de β sí y sólo sí $\mathbf{X}\overset{\wedge}{\beta} = \mathbf{P_X Y}$, donde $\mathbf{P_X}$ es el operador proyector ortogonal sobre $S(\mathbf{X})$, es decir, si y sólo si $\mathbf{X}\overset{\wedge}{\beta} = \mathbf{X}\left(\mathbf{X}^t\mathbf{X}\right)^{-}\mathbf{X}^t\mathbf{Y}$.*

Demostración: Por el Teorema 3.3.1 se cumple la relación

$$\begin{aligned}(\mathbf{Y}-\mathbf{X}\beta)^t(\mathbf{Y}-\mathbf{X}\beta) &= (\mathbf{Y}-\mathbf{P_X}\mathbf{Y})^t(\mathbf{Y}-\mathbf{P_X}\mathbf{Y}) \\ &+ (\mathbf{P_X}\mathbf{Y}-\mathbf{X}\beta)^t(\mathbf{P_X}\mathbf{Y}-\mathbf{X}\beta).\end{aligned}$$

Observe que ambos términos de la derecha son no negativos, y el primer término no depende de β. Por lo tanto $(\mathbf{Y}-\mathbf{X}\beta)^t(\mathbf{Y}-\mathbf{X}\beta)$ es minimizado cuando se minimiza

$$(\mathbf{P_X}\mathbf{Y}-\mathbf{X}\beta)^t(\mathbf{P_X}\mathbf{Y}-\mathbf{X}\beta),$$

la cual es mínima sí y sólo sí $\mathbf{P_X}\mathbf{Y}=\mathbf{X}\beta$, lo cual prueba el teorema. ■

Teorema 3.3.3 *β^* dado por $\beta^*=(\mathbf{X}^t\mathbf{X})^-\mathbf{X}^t\mathbf{Y}$ es un estimador de mínimos cuadrados de β.*

Demostración: Por el Teorema 3.3.2 se tiene que β^* es un estimador de mínimos cuadrados de β sí y sólo sí $\mathbf{X}\beta^*=\mathbf{P_X}\mathbf{Y}$, donde $\mathbf{P_X}$ es el operador proyector ortogonal sobre $S(\mathbf{X})$. En este caso se tiene que

$$\mathbf{X}\beta^*=\mathbf{X}\left[\left(\mathbf{X}^t\mathbf{X}\right)^-\mathbf{X}^t\mathbf{Y}\right]=\left[\mathbf{X}\left(\mathbf{X}^t\mathbf{X}\right)^-\mathbf{X}^t\right]\mathbf{Y}=\mathbf{P_X}\mathbf{Y}.$$

Por lo que, $\beta^*=(\mathbf{X}^t\mathbf{X})^-\mathbf{X}^t\mathbf{Y}$ es un estimador de mínimos cuadrados de β. ■

Otro parámetro de interés en el modelo lineal general es la varianza σ^2. Ahora se verá como el operador proyector ortogonal interviene en la estimación del parámetro σ^2.

Teorema 3.3.4 *Sea el modelo*

$$\mathbf{Y}=\mathbf{X}\beta+\mathbf{e}, \qquad E(\mathbf{e})=\mathbf{0}, \quad Cov(\mathbf{e})=\sigma^2\mathbf{I}. \tag{3.3}$$

Entonces

$$\frac{\mathbf{Y}^t(\mathbf{I}-\mathbf{P_X})\mathbf{Y}}{n-r(\mathbf{X})} \tag{3.4}$$

es un estimador insesgado de σ^2, respecto al modelo (3.2).

Demostración: Para la función cuadrática $\mathbf{Y}^t\mathbf{B}\mathbf{Y}$ se cumple la siguiente relación

$$E[\mathbf{Y}^t\mathbf{B}\mathbf{Y}] = tr(\mathbf{B}[Cov(\mathbf{Y})]) + [E(\mathbf{Y})]^t\mathbf{B}[E(\mathbf{Y})].$$

En el caso de interés se tiene que

$$\mathbf{B} = \mathbf{I} - \mathbf{P_X}, E(\mathbf{Y}) = \mathbf{X}\beta, Cov(\mathbf{Y}) = \sigma^2,$$

así se tiene que

$$E\left[\mathbf{Y}^t(\mathbf{I} - \mathbf{P_X})\mathbf{Y}\right] = tr(\sigma^2(\mathbf{I} - \mathbf{P_X})) + \beta^t\mathbf{X}^t(\mathbf{I} - \mathbf{P_X})\mathbf{X}\beta.$$

Además, se cumple que

$$tr(\sigma^2(\mathbf{I} - \mathbf{P_X})) = \sigma^2[tr(\mathbf{I} - \mathbf{P_X})] = \sigma^2(n - r\,(\mathbf{X}))$$

y $(\mathbf{I} - \mathbf{P_X})\mathbf{X} = \mathbf{0}$, así, $\beta^t\mathbf{X}^t(\mathbf{I} - \mathbf{P_X})\mathbf{X}\beta = \mathbf{0}$.

De lo anterior

$$E[\mathbf{Y}^t(\mathbf{I} - \mathbf{P_X})\mathbf{Y}] = \sigma^2(n - r\,(\mathbf{X})).$$

Así, se cumple que

$$E\left[\frac{\mathbf{Y}^t(\mathbf{I} - \mathbf{P_X})\mathbf{Y}}{n - r\,(\mathbf{X})}\right] = \sigma^2.$$

$\mathbf{Y}^t(\mathbf{I} - \mathbf{P_X})\mathbf{Y}$ se denomina la suma de cuadrados para el error (**SCE**). Al rango de $(\mathbf{I} - \mathbf{P_X})$ se le denomina los grados de libertad para el error (**gl**). Y a la razón entre la suma de cuadrados del error y los grados de libertad para el error se le denomina el error cuadrado medio (**ECM**).

Así, el error cuadrado medio respecto al modelo (3.2), está dado por medio de

$$\frac{\mathbf{Y}^t(\mathbf{I} - \mathbf{P_X})\mathbf{Y}}{n - r\,(\mathbf{X})}.$$

De lo anterior se tiene que una estimación insesgada de σ^2 está dada por la razón de la suma de cuadrados para el error dividida entre los grados de libertad para el error.

3.3.2. Mínimos cuadrados generalizados

Ahora considérese el modelo

$$\mathbf{Y} = \mathbf{X}\beta + \mathbf{e}, \qquad E(\mathbf{e}) = \mathbf{0}, \quad Cov(\mathbf{e}) = \sigma^2\mathbf{B}, \tag{3.5}$$

donde $\mathbf{B}$ es alguna matriz definida positiva. Así, existe una matriz no singular $\mathbf{R}$ para la cual se cumple que $\mathbf{B} = \mathbf{R}\mathbf{R}^t$. Del modelo (3.5) se obtiene el modelo

$$\mathbf{R}^{-1}\mathbf{Y} = \mathbf{R}^{-1}\mathbf{X}\beta + \mathbf{R}^{-1}\mathbf{e}, \quad E\left(\mathbf{R}^{-1}\mathbf{e}\right) = \mathbf{0}, \quad Cov\left(\mathbf{R}^{-1}\mathbf{e}\right) = \sigma^2\mathbf{I}. \tag{3.6}$$

Para el modelo (3.6) las estimaciones de mínimos cuadrados minimizan

$$\left(\mathbf{R}^{-1}\mathbf{Y} - \mathbf{R}^{-1}\mathbf{X}\beta\right)^t \left(\mathbf{R}^{-1}\mathbf{Y} - \mathbf{R}^{-1}\mathbf{X}\beta\right) = \left(\mathbf{Y} - \mathbf{X}\beta\right)^t \mathbf{V}^{-1} \left(\mathbf{Y} - \mathbf{X}\beta\right). \tag{3.7}$$

Esta distancia es una generalización de $(\mathbf{Y} - \mathbf{X}\beta)^t (\mathbf{Y} - \mathbf{X}\beta)$, por lo que los estimadores del parámetro β que minimizan tal distancia se conocen como estimadores por **mínimos cuadrados generalizados**.

Teorema 3.3.5 *Bajo el modelo (3.6), $\overset{\wedge}{\beta}$ es un estimador de mínimos cuadrados de β si y sólo si*

$$\mathbf{X}\overset{\wedge}{\beta} = \mathbf{X}\left(\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}^t\mathbf{V}^{-1}\mathbf{Y}. \tag{3.8}$$

En términos del operador proyector $\mathbf{P}_{\mathbf{XV}}$ del Teorema 3.3.2, el Teorema anterior se expresa, bajo el modelo (3.6), como: $\overset{\wedge}{\beta}$ es un estimador por mínimos cuadrados de β sí y sólo sí $\mathbf{X}\overset{\wedge}{\beta} = \mathbf{P}_{\mathbf{XV}}\mathbf{Y}$ donde

$$\mathbf{P}_{\mathbf{XV}} = \mathbf{X}\left(\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}^t\mathbf{V}^{-1} \tag{3.9}$$

es un proyector sobre $S(\mathbf{X})$.

Anteriormente se vió que una estimación insesgada de σ^2 está dada por la razón de **SCE** dividida entre los grados de libertad. Lo anterior bajo el modelo propuesto.

Bajo el modelo (3.2) la **SCE** estaba dada en términos del operador proyector ortogonal $\mathbf{P}_{\mathbf{X}}$ sobre $S(\mathbf{X})$ por medio de $\mathbf{Y}^t(\mathbf{I} - \mathbf{P}_{\mathbf{X}})\mathbf{Y}$.

En este caso, el modelo de interés es el dado en (3.6) por lo cual se tiene que la **SCE** está dada en términos del operador proyector ortogonal $\mathbf{P_R}$ sobre $S\left(\mathbf{R}^{-1}\mathbf{X}\right)$. El operador proyector ortogonal $\mathbf{P_R}$ sobre $S\left(\mathbf{R}^{-1}\mathbf{X}\right)$ está dado por medio de

$$\left(\mathbf{R}^{-1}\mathbf{X}\right)\left[(\mathbf{R}^{-1}\mathbf{X})^t(\mathbf{R}^{-1}\mathbf{X})\right]^-\left(\mathbf{R}^{-1}\mathbf{X}\right)^t.$$

Así, la **SCE** está dada por medio de $\left(\mathbf{R}^{-1}\mathbf{Y}\right)^t\left(\mathbf{I}-\mathbf{P_R}\right)\left(\mathbf{R}^{-1}\mathbf{Y}\right)$.

$$\mathbf{SCE}=\left(\mathbf{R}^{-1}\mathbf{Y}\right)^t\left[\mathbf{I}-\mathbf{R}^{-1}\mathbf{X}\left[(\mathbf{R}^{-1}\mathbf{X})^t(\mathbf{R}^{-1}\mathbf{X})\right]^-\left(\mathbf{R}^{-1}\mathbf{X}\right)^t\right]\left(\mathbf{R}^{-1}\mathbf{Y}\right).$$

Teorema 3.3.6 *Bajo el modelo (3.6), la* **SCE** *se puede expresar como*

$$\mathbf{Y}^t\left(\mathbf{I}-\mathbf{P_{XV}}\right)^t\mathbf{V}^{-1}\left(\mathbf{I}-\mathbf{P_{XV}}\right)\mathbf{Y}.$$

Demostración:

$$\begin{aligned}
\mathbf{SCE} &= \left(\mathbf{R}^{-1}\mathbf{Y}\right)^t\left[\mathbf{I}-\mathbf{R}^{-1}\mathbf{X}\left[(\mathbf{R}^{-1}\mathbf{X})^t(\mathbf{R}^{-1}\mathbf{X})\right]^-\left(\mathbf{R}^{-1}\mathbf{X}\right)^t\right]\left(\mathbf{R}^{-1}\mathbf{Y}\right)\\
&= \mathbf{Y}^t\mathbf{V}^{-1}\mathbf{Y}-\mathbf{Y}^t\mathbf{V}^{-1}\mathbf{X}\left(\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}\right)^-\mathbf{X}^t\mathbf{V}^{-1}\mathbf{Y}\\
&= \mathbf{Y}^t\left(\mathbf{I}-\mathbf{P_{XV}}\right)^t\mathbf{V}^{-1}\left(\mathbf{I}-\mathbf{P_{XV}}\right)\mathbf{Y}.
\end{aligned}$$

.

Se tiene que $\mathbf{R}$ es una matriz no singular, así se cumple que $r\left(\mathbf{R}^{-1}\mathbf{X}\right)=r\left(\mathbf{X}\right)$ y se tiene que los grados de libertad para el error bajo el modelo (3,6) están dados por medio de $n-r\left(\mathbf{X}\right)$.

De lo anterior se tiene que:

Teorema 3.3.7 *Bajo el modelo (3.6) una estimación insesgada para σ^2 está dada por medio de*

$$\frac{\mathbf{Y}^t\left(\mathbf{I}-\mathbf{P_{XV}}\right)^t\mathbf{V}^{-1}\left(\mathbf{I}-\mathbf{P_{XV}}\right)\mathbf{Y}}{n-r\left(\mathbf{X}\right)}. \tag{3.10}$$

Bibliografía

[1] Bapat y R. B. *Linear Algebra and Linear Models*, segunda edición, Springer-Verlag, Nueva York, EUA, 2000.

[2] Ben-Israel, A., y T. N. E. Greville, *Theory and aplications*, segunda edición, Springer-Verlag, Nueva York, EUA, 2008.

[3] Clarke, B.R. *Linear Models. The Theory and Application of analysis of variance* Wiley, Hoboken, NJ, EUA, 2008.

[4] Franklin A. Graybill, *Theory and Application of the Linear Model*, Wadsworth Brooks, EUA, 1976.

[5] Gentle, J. E., *Matrix Algebra. Theory, Computations, and Applications in Statistics*, Springer-Velarg, Nueva York, EUA, 2007.

[6] Kenneth Hoffman y Ray Kunze, *Álgebra Lineal*, segunda edición, Prentice Hall.

[7] Sengupta, D., y J. S. Rao, *Linear Models. An Integrated Approach*, World Scientific, Nueva Jersey, EUA, 2003.