Apéndice A

Álgebra lineal y vectores aleatorios

A.1. Vectores aleatorios

A.1.1. Definiciones

Sean $X \in R^p \quad Y \in R^q$ vectores aleatorios, es decir X un vector columna cuyos componentes son variables aleatorias, se define *valor esperado del vector* X al vector columna: $E(X) = [E(x_1), \ldots, E(x_p)]^T$ y matriz de covarianza entre X y Y:

$$Cov(X,Y) = \begin{bmatrix} cov(x_1,y_1) & .... & cov(x_1,y_q) \\ .... & .... & .... \\ cov(x_p,y_1) & .... & cov(x_p,y_q) \end{bmatrix} = E[(X - \mu x)(Y - \mu y)^T]$$

o su equivalente: $Cov(X,Y) = E[XY^T] - \mu x \, \mu y$
Nótese que $Cov(X,Y) = Cov(Y,X)^T$
Si $Y = X$ entonces se notará $Cov(X) = Cov(X,X)$.

A.1.2. Estimación

Si se dispone de una muestra de tamaño n de los vectores aleatorios X e Y donde la muestra de X está dada por la matriz $n \times p$, $Xobs$, y la muestra de Y dada por la matriz $n \times q$, $Yobs$, las medias muestrales mx y my se obtienen promediando las columnas de las matrices:

$$mx = \frac{1}{n} L_n^T Xobs \qquad my = \frac{1}{n} L_n^T Yobs$$

donde se nota al vector de n unos así: $L_n^T = [\ 1 \quad 1 \quad . \quad . \quad 1\ ]$.

En cuanto a la covarianza muestral está dada por:

$$Cov(X,Y) = \frac{1}{n}\, Xobs^T\, Yobs - \; mx\, my^T$$

A.1.3. Propiedades

Si A y B son matrices constantes: $Cov(AX, BY) = ACov(X,Y)B^T$
si además $X = Y$

- $Cov(AX, BX) = ACov(X)B^T$
- $Cov(AX) = ACov(X)A^T$

Si $a \in R^p$ y $b \in R^q$ son vectores columnas constantes entonces:

- $V(a^T X) = a^T Cov(X)a \geq 0$
- $cov(a^T X, b^T Y) = a^T Cov(X,Y)b$
- Si $X = Y$ entonces $cov(a^T X, b^T X) = a^T Cov(X)b$

Si A, B son matrices $m \times p$ y $m \times q$ constantes entonces:

$$Cov(AX+BY) = ACov(X)A^T + BCov(Y)B^T + ACov(X,Y)B^T + BCov(Y,X)A^T$$

Un caso particular es el siguiente: si $a \in R^p$ y $b \in R^q$ son vectores columnas constantes entonces:

$$Cov(a^T X + b^T Y) = a^T Cov(X)a + b^T Cov(Y)b + 2a^T Cov(X,\ Y)b$$

Si u es una variable aleatoria y $c \in R^n$, entonces:

$$Cov(u\, c) = Var(u)c\, c^T$$

Esta expresión se aplica a las proyecciones de un vector aleatorio $X \in R^p$ sobre la dirección de un vector c, fijo. Si se supone c unitario $\|c\| = 1$ la proyección está dada por $Px = (c^T X)c$ entonces:

$$Cov(Px) = c\, c^T Var(c^T X) = c\, c^T (c^T Cov(X)c)$$

nótese que el último paréntesis es un número real.
Si c es vector propio de $Cov(X)$ de valor propio λ, entonces: $Cov(Px) = c\, c^T\ \lambda c^T c$.
Y como $\|c\| = 1$ entonces $Cov(Px) = c\, c^T\ \lambda$.

A.2. Algunas fórmulas de derivación

A.2.1. Definición de gradiente

Si $x \in R^n \quad z \in R^q$ se define como gradiente del vector $z(x)$, ∇z la matriz $n \times q$ donde la columna k está formada por el vector gradiente $\nabla z_k(x)$ del k componente de z.

Regla de la cadena Si $y \in R^p$ es función de z, es decir $y(z)$ entonces el gradiente de la función compuesta $y(z(x))$, respecto de x, está dado por:

$$\nabla_x y = \nabla_x z \times \nabla_z y$$

nótese que el producto es conforme: $n \times p = (n \times q)(q \times p)$.

Ejemplo: Si la matriz $A \quad (m \times q)$ es constante entonces $\nabla Ax = A^T$; luego, si $y(x) \in R^q$ es función de x , por la regla de la cadena:
$\nabla Ay(x) = \nabla_x y(x)\, A^T$.

A.2.2. propiedades:

Si $\quad y(x) \in R^q$ entonces $\nabla(z^T y) = \nabla z\; y + \nabla y\; z$

A.2.2.1. Corolarios

- si $z = y$ entonces $\nabla(\|y\|^2) = 2\nabla y\; y$
- $\nabla(\|x\|^2) = 2x$
- si $z = constante$ entonces $\nabla(z^T y) = \nabla y\; z$
- $\nabla(z^T x) = \; z$
- $\nabla \|Ay\|^2 = 2\nabla y A^T A y$
- $\nabla \|Ax\|^2 = 2A^T Ax$
- $Hessiano(\|Ax\|^2) = 2A^T A$
- $\nabla(x^T Ax) = (A^T + A)x$, si A es simétrica: $\nabla(x^T Ax) = 2Ax$
- $Hessiano\,(x^T Ax) = A^T + A$, si A es simétrica: $Hessiano\,(x^T Ax) = 2A$

A.3. Solución de un sistema frecuente en Kriging

Sea el sistema cuadrado ya particionado de $n + m$ ecuaciones, $n > m$, de solución $\begin{pmatrix} \alpha \\ \lambda \end{pmatrix}$:

$$\begin{array}{ccc} C & F & = & e \\ F^T & 0 & = & g \end{array}$$

con C $(n*n)$ definida positiva, F $n*m$ de rango m.

Este sistema tiene solución única, la matríz es invertible; si no lo fuera existiría una solución del sistema homogéneo:

$$\begin{array}{c} C\alpha + F\lambda = 0 \\ F^T\alpha = 0 \end{array}$$

con al menos uno de los dos, $\alpha \neq 0 \quad o \ \lambda \neq 0$. Entonces por la primera ecuación si $\alpha = 0 \Rightarrow F\lambda = 0$ pero como F es de rango m , $\lambda = 0$.

Por otra parte si $\alpha \neq 0$, multiplicando la primera por α^T queda $\alpha^T C\alpha + \alpha^T F\lambda = 0$, pero por la segunda $\alpha^T F\lambda = 0$; luego, $\alpha^T C\alpha = 0$ que no puede ser ya que C es definida positiva.

Habiéndose probado la existencia, resolvamos el sistema. Multiplicando la primera por C^{-1} queda:

$$\begin{array}{ccccc} I & C^{-1}F & = & C^{-1}e \\ F^T & 0 & = & g \end{array}$$

Multiplicando la primera por $-F^T$ y sumando a la segunda queda:

$$\begin{array}{cccc} I & C^{-1}F & = & C^{-1}e \\ 0 & -F^T\, C^{-1}F & = & g - F^T C^{-1}e \end{array}$$

Luego, la solución del sistema está dada por

$$\lambda = inv(F^T\ C^{-1}F)\ (F^T C^{-1}e - g)$$

y

$$\alpha = C^{-1}e - C^{-1}F\,\lambda$$

$$\left\{ \begin{array}{ccc} \alpha & = & q \\ q & = & q \\ q & = & q \end{array} \right\}$$

Si $F = L$ entonces $\lambda = \frac{L^T C^{-1}e - g}{L^T\ C^{-1}L}$ es un escalar. En cuanto a $\alpha = C^{-1}e - C^{-1}\lambda L$.