

APÉNDICE F

F. MÍNIMOS CUADRADOS

Sea \mathbf{a} un vector de pesos e y_i un vector con d componentes, de forma que los productos internos $\mathbf{a}^t y_i$ son positivos, se trata de hacer $\mathbf{a}^t y_i = b_i$, donde los b_i son constantes positivas. El objetivo es resolver un conjunto de ecuaciones lineales. Para ello se introduce la notación matricial de forma que sea Y una matriz de dimensiones n -por- d cuya i -ésima fila es el vector y_i^t y sea \mathbf{b} el vector columna $\mathbf{b} = (b_1, \dots, b_n)^t$. El problema es encontrar un vector \mathbf{a} que satisfaga $Y\mathbf{a} = \mathbf{b}$. Si Y fuera no singular, se podría escribir $\mathbf{a} = Y^{-1}\mathbf{b}$ y obtener una solución formal. No obstante, Y es rectangular, generalmente con más filas que columnas, en cuyo caso \mathbf{a} está sobredeterminado, y ordinariamente no existe una solución formal. No obstante, se puede buscar un vector \mathbf{a} que minimice alguna función de error entre $Y\mathbf{a}$ y \mathbf{b} . Si se define el vector error \mathbf{e} por

$$\mathbf{e} = Y\mathbf{a} - \mathbf{b} \quad (\text{F.1})$$

entonces una aproximación consiste en tratar de minimizar la suma de los errores al cuadrado, que es equivalente a minimizar la función criterio suma de los errores al cuadrado

$$J_s(\mathbf{a}) = \|Y\mathbf{a} - \mathbf{b}\|^2 = \sum_{i=1}^n (\mathbf{a}^t y_i - b_i)^2 \quad (\text{F.2})$$

El problema de minimizar la suma de los errores al cuadrado es un problema clásico, que puede resolverse por el procedimiento del gradiente,

$$\nabla J_s = \sum_{i=1}^n 2(\mathbf{a}^t \mathbf{y}_i - b_i) \mathbf{y}_i = 2Y^t(Y\mathbf{a} - \mathbf{b}) \quad (\text{F.3})$$

e igualándolo a cero se obtiene la condición necesaria,

$$Y^t Y \mathbf{a} = Y^t \mathbf{b} \quad (\text{F.4})$$

de esta forma el problema de resolver $Y\mathbf{a} = \mathbf{b}$ se convierte en el de resolver $Y^t Y \mathbf{a} = Y^t \mathbf{b}$. Esta última ecuación tiene la gran ventaja de que la matriz *d-por-d* $Y^t Y$ es cuadrada y muy a menudo no singular. Si es no singular, se puede obtener \mathbf{a} de forma unívoca,

$$\mathbf{a} = (Y^t Y)^{-1} Y^t \mathbf{b} = Y^t \mathbf{b} \quad (\text{F.5})$$

donde la matriz *d-por-n* Y^t se denomina la pseudoinversa de Y . Esta matriz tiene la particularidad de que si Y es invertible entonces $Y^t = Y^{-1}$.

Cuando las componentes de los vectores y los elementos de las matrices dados en (F.1) y (F.2) son valores numéricos complejos, la expresión (F.3) resulta ser,

$$\nabla J_s = 2Y^*(Y\mathbf{a} - \mathbf{b}) \quad (\text{F.6})$$

donde Y^* es la matriz traspuesta conjugada de Y , en cuyo caso la ecuación (F.5) se convierte en,

$$\mathbf{a} = (Y^* Y)^{-1} Y^* \mathbf{b} = Y^\# \mathbf{b} \quad (\text{F.7})$$

donde la matriz $Y^\# = (Y^* Y)^{-1} Y^*$ se le denomina a menudo como matriz *inversa generalizada de Moore-Penrose* (Noble 1969).