PESQUIMAT, Revista de la F.C.M. de la
Universidad Nacional Mayor de San Marcos
Vol. XII N°2, pp. 74-85, Lima - Perú, Agosto 2009

Minimización de Cuadráticas

Edinson R. Montoro Alegre[1], *Martha Hilda Timoteo Sanchez*[2]

Resumen: En este trabajo, presentamos el problema de Minimización de Cuadráticas sin restricciones y se da una técnica de solución para cada uno de los diferentes casos que se pueden presentar.
Palabras clave: Programación matemática, programación convexa, diagonalización, descomposición espectral y factorización de Cholesky.

QUADRATIC MINIMIZATION

Abstract: In this work, we present the Quadratic Minimization problems without restrictions and it's given a solution technique for each of the different cases which can be presented.
Key words: Mathematic programation, convex programation, diagonalization, spectral decomposition and Cholesky decomposition.

1. Introducción

Una cuadrática es un polinomio en n variables con términos hasta segundo orden. La minimización de estas funciones tienen interés por el gran número de aplicaciones que recaen en este formato [3], [4] y [5]. Por ejemplo, cuando para un conjunto de datos impíricos se postula una relación linear con ciertos parámetros desconocidos, el problema de ajustar esos parámetros suele ser resuelto a través de la minimización de la suma de los cuadrados de los errores, es decir, la minimización de una función cuadrática. La suma de cuadrados no es mejor que otras medidas globales de errores, en términos de calidad de ajuste (pues se pudo haber considerado él con otro tipo de norma). Sin embargo, es la medida cuya minimización es la más simple del punto de vista numérico. De hecho la minimización de cuadráticas es uno de los problemas más fáciles en el arte de la OPTIMIZACIÓN, haciendo también que sea utilizado frecuentemente como *subproblema auxiliar* en algoritmos que resuelven problemas más complicados.

2. Cuadráticas sin Restricciones

Dada $G \in \mathbb{R}^{n\times n}$ simétrica, $b \in \mathbb{R}^n$ y $c \in \mathbb{R}$ constante. El problema de estudio será

$$\text{minimizar } q(x) = \frac{1}{2}x^t Gx + b^t x + c \tag{1}$$

Lema 2.1.
Si $q(x) = \frac{1}{2}x^t Gx + b^t x + c$ *entonces*

$$\nabla q(x) = Gx + b$$
$$Hq(x) = G$$

[1] Universidad Nacional Mayor de San Marcos - Facultad de Ciencias Matemáticas, e-mail: edinsonmontoro@yahoo.com

[2] Universidad Nacional Mayor de San Marcos - Facultad de Ciencias Matemáticas, e-mail: mtimoteos@yahoo.com

Prueba.
Sea
$$x = (x_1, x_2, \ldots, x_n), \qquad b = (b_1, b_2, \ldots, b_n)$$
y
$$G = \begin{Bmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ g_{n1} & g_{n2} & \cdots & g_{nn} \end{Bmatrix}$$
entonces
$$q(x) = \frac{1}{2}\left[x'\left(\sum_{i=1}^{n} g_{1i}x_i, \sum_{i=1}^{n} g_{2i}x_i, \ldots, \sum_{i=1}^{n} g_{ni}x_i\right)\right] + \left(\sum_{i=1}^{n} b_i x_i\right) + c$$
$$q(x) = \frac{1}{2}\left[\sum_{i=1}^{n} g_{1i}x_i x_1 + \sum_{i=1}^{n} g_{2i}x_i x_2 + \ldots + \sum_{i=1}^{n} g_{ni}x_i x_n\right] + \sum_{i=1}^{n} b_i x_i + c.$$
Luego
$$\nabla q(x) = \left(\frac{\partial q(x)}{\partial x_1}, \frac{\partial q(x)}{\partial x_2}, \ldots, \frac{\partial q(x)}{\partial x_n}\right)$$
es decir
$$\frac{\partial q(x)}{\partial x_1} = \frac{1}{2}\left[\left(g_{11}x_1^2 + \sum_{i=2}^{n} g_{1i}x_i x_1\right)' + (g_{21}x_1x_2)' + \ldots + (g_{n1}x_1x_n)'\right] + b_1$$
$$\frac{\partial q(x)}{\partial x_1} = \frac{1}{2}\left[2g_{11}x_1 + \sum_{i=2}^{n} g_{1i}x_i + (g_{21}x_2 + g_{31}x_3 + \ldots + g_{n1}x_n)\right] + b_1$$
pero por ser G simétrica se cumple que $g_{ij} = g_{ji}$. Por tanto se obtiene
$$\frac{\partial q(x)}{\partial x_1} = g_{11}x_1 + \frac{1}{2}\left[\sum_{i=2}^{n} 2g_{1i}x_i\right] + b_1$$
$$\frac{\partial q(x)}{\partial x_1} = g_{11}x_1 + \sum_{i=2}^{n} g_{1i}x_i + b_1 = \sum_{i=1}^{n} g_{1i}x_i + b_1.$$
En forma análoga se obtiene un resultado similar para cada
$$\frac{\partial q(x)}{\partial x_j} = \sum_{i=1}^{n} g_{ji}x_i + b_j \qquad j = 2, 3, \ldots, n$$
Luego
$$\nabla q(x) = \left(\frac{\partial q(x)}{\partial x_1}, \frac{\partial q(x)}{\partial x_2}, \ldots, \frac{\partial q(x)}{\partial x_n}\right) = \left(\sum_{i=1}^{n} g_{1i}x_i + b_1, \ldots, \sum_{i=1}^{n} g_{ni}x_i + b_n\right)$$
$$\nabla q(x) = \left(\sum_{i=1}^{n} g_{ji}x_i, \ldots, \sum_{i=1}^{n} g_{ji}x_i\right) + (b_1, \ldots, b_n)$$
$$\nabla q(x) = \begin{pmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ g_{n1} & g_{n2} & \cdots & g_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

$$\nabla q(x) = Gx + b.$$

Para la última parte del lema hacemos

$$h(x) = (h_1(x), \ldots, h_n(x)) = \nabla q(x) = \left(\frac{\partial q(x)}{\partial x_1}, \frac{\partial q(x)}{\partial x_2}, \ldots, \frac{\partial q(x)}{\partial x_n}\right)$$

$$\nabla h(x) = \begin{pmatrix} \frac{\partial h_1(x)}{\partial x_1} & \frac{\partial h_1(x)}{\partial x_2} & \cdots & \frac{\partial h_1(x)}{\partial x_n} \\ \frac{\partial h_2(x)}{\partial x_1} & \frac{\partial h_2(x)}{\partial x_2} & \cdots & \frac{\partial h_2(x)}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial h_n(x)}{\partial x_1} & \frac{\partial h_n(x)}{\partial x_2} & \cdots & \frac{\partial h_n(x)}{\partial x_n} \end{pmatrix} = \begin{pmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ g_{n1} & g_{n2} & \cdots & g_{nn} \end{pmatrix} = G$$

■

Los puntos críticos de (1) serán aquellos puntos donde el gradiente se anula, por tanto de acuerdo al lema 1, serán todas las soluciones del sistema lineal

$$Gx + b = 0 \tag{2}$$

La existencia o unicidad estará determinada por las propiedades de este sistema.

Lema 2.2.

a) ***El problema (1) admite algún punto crítico si y solo si*** $b \in \mathcal{R}(G)$, *($\mathcal{R}(G)$ es el espacio generado por las columnas de G).*

b) ***El problema (1) admite un único punto crítico si y solo si*** G *es no singular.*

Prueba.

a) Sea x^* un punto crítico, entonces

$$Gx^* = -b$$

es decir

$$[G_1, G_2, \ldots, G_n] \begin{bmatrix} x_1^* \\ x_2^* \\ \vdots \\ x_n^* \end{bmatrix} = -b$$

donde G_i representa la i-ésima columna de G. De donde

$$x_1^* G_1 + x_2^* G_2 + \ldots + x_n^* G_n = -b.$$

Por lo tanto

$$b \in \mathcal{R}(G)$$

La parte recíproca es análoga.

b) Tenemos el sistema lineal

$$Gx + b = 0$$

ó

$$Gx = -b \tag{$*$}$$

el cual es un sistema de n ecuaciones y n incógnitas. Por el álgebra lineal $(*)$ posee solución única si y solo si G es una matriz no singular.

■

La ecuación de los puntos críticos

$$Gx + b = 0 \tag{2}$$

puede tener una, infinita o ninguna solución.

Si (2) no tiene solución, es decir $b \notin \mathcal{R}(G)$ entonces, (1) no admite minimizador local o global (caso contrario se tendría un x^* tal que $Gx^* + b = 0$. Contradicción con la hipótesis). Uno de esos casos, podría ser por ejemplo cuando

$$q(x) = \frac{1}{2}x^t Gx + b^t x + c$$

con $G = 0$ y $b \neq 0$.

Si (2) tiene solución única, entonces ese punto sería el único punto crítico de (1), pero él podría ser un minimizador, maximizador o punto silla.

Finalmente si (2) tiene infinitas soluciones (eso ocurre cuando G es singular y $b \in \mathcal{R}(G)$) entonces todas ellas serán puntos críticos y del mismo tipo (será probado más adelante).

Observación:

Es interesante observar aquí, que en un problema con infinitas soluciones (G singular y $b \in \mathcal{R}(G)$) puede ser transformado en un problema sin solución, debido a una perturbación pequeña en el vector b. Por ejemplo el sistema lineal

$$0x + 0 = 0$$

tiene como soluciones a todo $\mathbb{R}^n$, pero el sistema

$$0x + \varepsilon = 0$$

no posee solución alguna para cualquier $\varepsilon \neq 0$.

Esto nos muestra que, muchas veces, es difícil distinguir las situaciones "sin solución" de las "infinitas soluciones". En efecto: Debido a errores de redondeo, puede ocurrir que, el vector b que, en realidad estaba en el espacio columna de G ($b \in \mathcal{R}(G)$), quede fuera de ese subespacio, haciendo que un sistema con infinitas soluciones aparente ser incompatible en los cálculos numéricos.

También puede ocurrir que una matriz G singular se torne *inversible*, por perturbaciones de redondeo (las clásicas matrices mal-condicionadas), transformando un sistema incompatible, o indeterminado en un problema con solución única.

Usando resultados de convexidad y las condiciones de optimilidad de segundo orden (ver [1]), podemos clasificar fácilmente los puntos críticos de (1):

Si x^* es minimizador local, entonces, se cumple $Hq(x^*) \geq 0$. Pero $G = Hq(x^*)$, entonces G es semidefinida positiva para todo x. Luego $q(x)$ será una función convexa y cualquier minimizador local es global. Con todo esto podemos concluir: *Si x^* es un punto crítico y $G \geq 0$ $\forall x \in \mathbb{R}^n$ entonces necesariamente x^* es un minimizador global.*

Con el mismo razonamiento deducimos que toda cuadrática tiene un único tipo de punto crítico, por que la hessiana es una matriz constante, es decir todos serán minimizadores globales, maximizadores globales o puntos sillas.

La prueba del siguiente lema, muestra que, debido a la simplicidad de las funciones cuadráticas, es fácil obtener las mismas conclusiones sin apelar a los resultados de convexidad.

Lema 2.3.

Si $G \geq 0$ y x^* punto crítico de (1), entonces x^* es un minimizador global de (1).

Prueba.

Tenemos

$$q(x) = \frac{1}{2}x^t Gx + b^t x + c = \frac{1}{2}x^t Gx - x^* Gx + c$$

pues $x^*G = -b$. Luego

$$q(x) = \frac{1}{2}x^tGx + \frac{1}{2}x^*Gx^* - x^*Gx - \frac{1}{2}x^*Gx^* + c$$

$$q(x) = \frac{1}{2}(x - x^*)^tG(x - x^*) - \frac{1}{2}x^*Gx^* + c$$

Luego por ser G semidefinida positiva se tiene

$$q(x) \geq -\frac{1}{2}x^*Gx^* + c = -\frac{1}{2}x^*Gx^* + \frac{1}{2}x^*Gx^* - \frac{1}{2}x^*Gx^* + c \qquad \forall x \in \mathbb{R}^n$$

$$q(x) \geq \frac{1}{2}x^*Gx^* - x^*Gx^* + c = \frac{1}{2}x^*Gx^* + b^tx^* + c = q(x^*)$$

$$q(x) \geq q(x^*) \qquad \forall x \in \mathbb{R}^n$$

Por lo tanto x^* es un minimizador global. ■

Lema 2.4.
Si (1) admite un minimizador local, entonces G es semidefinida positiva para todo $x \in \mathbb{R}^n$.

Prueba.
Si x^* es un minimizador local, entonces $\exists r > 0$ tal que

$$q(x) \geq q(x^*) \qquad \forall x \in B(x^*, r)$$

$$\frac{1}{2}xGx + b^tx + c \geq \frac{1}{2}x^*Gx^* + b^tx^* + c$$

de donde

$$\frac{1}{2}xGx - \frac{1}{2}x^*Gx^* + b^tx - b^tx^* \geq 0$$

$$\frac{1}{2}xGx - \frac{1}{2}x^*Gx^* + b^tG(x - x^*) \geq 0$$

$$\frac{1}{2}xGx - \frac{1}{2}x^*Gx^* - x^*G(x - x^*) \geq 0$$

$$\frac{1}{2}xGx - \frac{1}{2}x^*Gx^* - x^*Gx + x^*Gx^* \geq 0$$

$$\frac{1}{2}xGx - x^*Gx + \frac{1}{2}x^*Gx^* \geq 0$$

$$\frac{1}{2}(x - x^*)^tG(x - x^*) \geq 0 \qquad \forall x \in B(x^*, r) \tag{α}$$

Sea $x \notin B(x^*, r)$, entonces

$$y = \frac{x - x^*}{\|x - x^*\|}a; \qquad \text{con } a < r$$

es tal que $y \in B(x^*, r)$. Luego

$$0 \leq \left(\frac{x - x^*}{\|x - x^*\|}a\right) G \left(\frac{x - x^*}{\|x - x^*\|}a\right) = \frac{a^2}{\|x - x^*\|^2}\left[(x - x^*)^tG(x - x^*)\right]$$

de donde

$$(x - x^*)^tG(x - x^*) \geq 0 \qquad \forall x \notin B(x^*, r) \tag{β}$$

De (α) y (β) concluimos que

$$G \geq 0 \qquad \forall x \in \mathbb{R}^n$$

■

Corolario 2.1.
Todo minimizador local de (1) es global.

Prueba.
Si x^* es minimizador local, entonces por lema 4 $G \geq 0$ $\forall x \in \mathbb{R}^n$.
Por otro lado, por teorema de Taylor se tiene que

$$q(x) = q(x^*) + \nabla q(x^*)(x - x^*) + \frac{1}{2}(x - x^*)Hq(z)(x - x^*) \qquad \forall x$$

$$q(x) = q(x^*) + [Gx^* + b](x - x^*) + \frac{1}{2}(x - x^*)G(x - x^*) \qquad \forall x$$

$$q(x) = q(x^*) + 0 + \frac{1}{2}(x - x^*)G(x - x^*) \qquad \forall x$$

$$q(x) \geq q(x^*) \qquad \forall x \in \mathbb{R}^n$$

Por lo tanto x^* es minimizador global. ■

Corolario 2.2.
Si la matriz G es indefinida, entonces la cuadrática $q(x)$ no tiene extremos locales (es decir no posee ni máximos ni mínimos locales).

Prueba.
Sea x^* un punto crítico.
Si G es indefinida, entonces $G = Hq(x^*)$ es indefinida. Luego, existirán w, y vectores no nulos en $\mathbb{R}^n$ tal que

$$y.H(x^*)y > 0 \qquad w.H(x^*)w < 0$$

Como $q(x)$ tiene segundas derivadas parciales continuas, existirá un $\varepsilon > 0$ tal que

$$y.H(x^* + ty)y > 0 \qquad w.H(x^* + tw)w < 0$$

para todo t tal que $|t| < \varepsilon$. Entonces, podemos definir

$$Y(t) = q(x^* + ty) \qquad W(t) = q(x^* + tw),$$

de donde

$$Y'(0) = 0 = W'(0)$$

y

$$Y''(0) = y.Hq(x^*)y > 0 \qquad W''(0) = w.Hq(x^*)w < 0.$$

Por tanto, podemos concluir que, $t = 0$ mínimo local estricto para $Y(t)$ y máximo local estricto para $W(t)$. Esto quiere decir, que si nos movemos a partir de x^* en la dirección de y ó $-y$, los valores de $q(x)$ aumentan, pero si nos movemos a partir de x^* en la dirección de w ó $-w$, los valores de $q(x)$ disminuyen. Por esta razón, denominamos al punto crítico x^* *punto silla*, y este punto no es ni máximo ni mínimo. ■

3. Solución de Cuadráticas sin Restricciones usando factorización

La forma más ruda de resolver (1) es considerando la descomposición espectral de G. En efecto, como G es simétrica, existe una matriz ortogonal Q (es decir $QQ^t = Q^tQ = I$) y una matriz diagonal D tal que

$$G = QDQ^t. \tag{3}$$

Los autovalores de G, $\sigma_1, \sigma_2, \ldots, \sigma_n$, son elementos de la diagonal D y los autovalores correspondientes son las columnas de Q. Luego para cualquier $x \in \mathbb{R}^n$, $x \neq 0$, existirá un $y \in \mathbb{R}^n$, $y \neq 0$ tal que

$$x = Qy = [q_1]y_1 + [q_2]y_2 + \ldots + [q_n]y_n$$

donde los $[q_i]$ son las columnas de Q. Entonces

$$x^tGx = (Qy)^tG(Qy) = y^t(Q^tGQ)y$$

$$x^tGx = y^tDy = \sigma_1 y_1^2 + \sigma_2 y_2^2 + \ldots + \sigma_n y_n^2$$

De donde,

$$x^tGx \geq 0 \Longleftrightarrow \sigma_i \geq 0 \; \forall i.$$

Con eso afirmamos: G es semidefinida positiva si todas las entradas de D son no negativas. Si todos los elementos de la diagonal son mayores que cero, entonces, D y G son definidas positivas.

Por tanto, observando la diagonal de D obtenemos información sobre el tipo de puntos críticos que el problema (1) puede tener. Si estamos interesados en minimizadores, y $D \geq 0$, entonces analizaremos el sistema

$$Gx + b = 0.$$

Usando (3), el sistema toma la forma

$$(QDQ^t)x = -b \tag{4}$$

De la cual obtenemos:

$$DQ^tx = -Q^tb \tag{5}$$

$$Dz = -Q^tb \tag{6}$$

donde $z = Q^tx$. Además

$$x = Qz \tag{$*$}$$

El sistema (6) tendrá solución si y solo si, un posible cero en la diagonal de D corresponde a una coordenada nula ($[Q^tb]_i$) de $-Q^tb$.

Si hay un cero en la diagonal, es decir $\sigma_i = 0$, tal que $[Q^tb]_i \neq 0$, entonces, el sistema (4) no posee solución, y consecuentemente el problema (1) carece de puntos críticos.

Si todos los elementos de D son estrictamente positivos (i.e. $\det(G) \neq 0$), entonces, (4) tiene solución única, y el vector x calculado a través de (6) y ($*$) será un minimizador global del problema (1).

Si el sistema (6) es compatible, pero existe $\sigma_i = 0$ y $[Q^tb]_i = 0$, entonces tendremos infinitas soluciones. En efecto, sea

$$Dz = -Q^tb$$

$$\begin{pmatrix} \sigma_1 z_1 \\ \vdots \\ \sigma_i z_i \\ \vdots \\ \sigma_j z_j \\ \vdots \\ \sigma_n z_n \end{pmatrix} = \begin{pmatrix} [-Q^t b]_1 \\ \vdots \\ [-Q^t b]_i \\ \vdots \\ [-Q^t b]_j \\ \vdots \\ [-Q^t b]_n \end{pmatrix}$$

con $\sigma_i = 0$ y $[-Q^t b]_i = 0$; $\sigma_j = 0$ y $[-Q^t b]_j = 0$. Notamos que z_i podrá tomar cualquier valor y siempre será solución. Lo mismo ocurre para z_j. Esto quiere decir. que se obtendrán infinitas soluciones. Cabe mencionar aquí, que debido a que

$$z = Q^t x \text{ entonces } x = Qz$$

es decir

$$x = [q_1 \; q_2 \; \dots \; q_n] z$$

donde q_i son las soluciones de Q. Entonces

$$x = z_1 q_1 + z_2 q_2 + \dots + z_n q_n.$$

Pero, como z_i y z_j toman infinitos valores, x tomará la forma

$$x = M + z_i q_i + z_j q_j$$

es decir

$$x \in A = \{x = M + z_i q_i + z_j q_j / \; z_i, \; z_j \in \mathbb{R}\}$$

Esto quiere decir que, todas las infinitas soluciones de (4) forman una variedad afín en $\mathbb{R}^n$ de dimensión igual al número de ceros en la diagonal D. Además todas las soluciones serán minimizadores, pues por dato $D \geq 0$ lo que implica $G \geq 0$.

Analicemos ahora los diferentes casos que se pueden presentar:

Caso 1. Cuando no existen minimizadores (es decir G es indefinida) del problema (1), es útil determinar para cada $x \in \mathbb{R}^n$ una dirección $d \in \mathbb{R}^n$ tal que

$$\lim_{t \to \infty} q(x + td) = -\infty \tag{7}$$

Si realmente supiéramos hallar una dirección que satisfaga (7), podríamos decir que siempre seremos capaces de resolver (1) (incluso cuando el mínimo sea $-\infty$, minimizador será $x + \infty d$).

Analicemos pues este caso:

Si algún autovalor de G, $\sigma_i < 0$, entonces tomamos d =autovector asociado a σ_i (ubicada en la $i-$ésima columna de Q), entonces

$$\begin{aligned} q(x + td) &= \frac{1}{2}(x + td)^t G(x + td) + b^t(x + td) + c \\ &= \frac{1}{2} x^t G x + x^t G(td) + \frac{1}{2} t^2 d^t G d + b^t x + t b^t d + c \\ &= q(x) + t(b^t d + x^t G d) + \frac{1}{2} t^2 d^t G d \\ &= q(x) + t(b^t + Gx)d + \frac{1}{2} t^2 d^t G d \\ &= q(x) + t \nabla q(x) d + \frac{1}{2} t^2 d^t G d \end{aligned}$$

Pero como $d^tGd = d^t(QDQ^t)d = (d^tQ)D(Q^td) = (0 \ \ldots \ 1 \ \ldots \ 0)D \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \sigma_i$

Por lo tanto

$$q(x + td) = q(x) + t\nabla q(x)d + \frac{1}{2}t^2\sigma_i$$

Con esto podremos afirmar que $q(x + td)$ como función de t es una parábola cóncava y

$$q(x + td) \rightarrow -\infty \quad \text{cuando} \quad t \rightarrow \pm\infty.$$

La elección de aquel d, no es la única que satisface (7). En efecto bastará escoger cualquier dirección d tal que $d^tGd < 0$.

Caso 2. Consideremos el caso en que $D \geq 0$, pero existe $\sigma_i = 0$ con $[Q^tb]_i \neq 0$ (es decir el sistema $Gx + b = 0$ es incompatible, no existe solución alguna). Tomemos nuevamente d =autovector asociado a σ_i

$$b^td = \left[Q^tb\right]_i \neq 0 \quad \text{y} \quad d^tGd = \sigma_i = 0$$

siempre que $b^td = [Q^tb]_i < 0$, caso contrario se tomará $-d$. De esta manera, siempre podremos suponer que $b^td < 0$. Luego

$$\begin{aligned} q(x + td) =& q(x) + t\nabla q(x)d + \frac{1}{2}d^tGd \\ =& q(x) + t(Gx + b)^td \\ =& q(x) + tx^tGd + tb^td \end{aligned}$$

Pero $Gd = 0$. En efecto

$$\begin{aligned} Gd =& (QDQ^t)d = (QD)(Q^td) = (QD) \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \\ =& Q\left[\begin{pmatrix} \sigma_1 & & & & \\ & \ddots & & & \\ & & \sigma_i & & \\ & & & \ddots & \\ & & & & \sigma_n \end{pmatrix}\begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}\right] = Q\begin{pmatrix} 0 \\ \vdots \\ \sigma_i \\ \vdots \\ 0 \end{pmatrix} = 0 \end{aligned}$$

por lo tanto

$$q(x + td) = q(x) + tb^td$$

Luego, $q(x + td)$ es una recta con pendiente negativa y

$$q(x + td) \rightarrow -\infty \quad \text{cuando} \quad t \rightarrow \infty.$$

Con todo esto podemos concluir que, la Descomposición Espectral, resuelve de manera totalmente satisfactoria el problema (1). Sin embargo, su costo computacional es frecuentemente intolerable, y la búsqueda de alternativas más baratas es necesaria.

La manera más popular de resolver (1), se basa en la *Factorización de Cholesky* de G (ver [2]). Tal procedimiento funciona y es estable solo cuando G es definida positiva, en tal caso

$$G = LDL^t$$

donde $L \in \mathbb{R}^{n\times n}$ matriz triangular inferior y elementos en la diagonal iguales a 1, $D \in \mathbb{R}^{n\times n}$ matriz diagonal con elementos positivos.

El siguiente algoritmo halla los factores Cholesky de G.

Algoritmo

$d_{11} := g_{11}$

Para j=2 **hasta** n **hacer**

$d_{jj} := g_{11} - \sum\limits_{k=1}^{j-1} d_{kk} l_{jk}^2$

Si $j = n$ **Pare.**

Si $j < n$ **Entonces**

Para $i = j + 1$ **hasta** n **hacer**

$l_{ij} := \dfrac{1}{d_{jj}} \left(g_{ij} - \sum\limits_{k=1}^{j-1} d_{kk} l_{jk} l_{ik} \right)$

El algoritmo de Cholesky termina produciendo $D > 0 \iff G$ es definida positiva. Si G es singular o indefinida, en algún momento aparecerá un $d_{jj} \leq 0$ en el cálculo de estas entradas.

En los casos en que la factorización de Cholesky de G es completada con éxito, el único minimizador del problema (1) es obtenido resolviendo

$$LDL^t x = -b$$

proceso que puede ser descompuesto en tres pasos:

a) Resolver $Ly = -b$

b) Resolver $Dz = y$

c) Resolver $L^t x = z$

La minimización de la función cuadrática bajo este procedimiento alcanza $O(n^3/6)$ de tiempo.

Cuando, con el algoritmo, detectamos que la matriz G no es definida positiva, podemos apelar por el proceso mucho más costoso de calcular su descomposición espectral.

4. Aplicaciones

4.1. Selección de Cartera

Un inversionista tiene P dólares para invertir en un grupo de n acciones y quisiera saber cuanto invertir en cada una de ellas. La combinación seleccionada se conoce como *cartera del inversionista*. El inversionista tiene metas contradictorias: le gustaría tener una cartera que tuviera al mismo tiempo un gran rendimiento esperado y un riesgo pequeño. Estas metas son contradictorias porque lo más frecuente es que, en el mundo real, las carteras con alto rendimiento esperado tienen también alto riesgo.

Supóngase que una inversión de D_i dólares se aplica al activo i y suponga que durante algún periodo de tiempo especificado estos D_i dólares se convierten en $1.3D_i$. En este caso diríamos que el **rendimiento** durante ese periodo es de

$$\frac{1.3D_i - D_i}{D_i} = 0.3$$

Por fines prácticos se supondrá que el riesgo se mide **mediante la varianza del rendimiento en la cartera.**

Formularemos por escrito el modelo general para un problema de dos activos. Se usará la siguiente notación:
σ_i^2 =Varianza de los rendimientos anuales de la acción i. $i = 1, 2$
σ_{12} =Covarianza de los rendimientos anuales de las acciones 1 y 2.
R_i =Rendimiento anual esperado de la acción i. $i = 1, 2$
G =Límite inferior sobre el rendimiento anual esperado de la inversión total.
S_i =Límite superior sobre la inversión en la acción i, $i = 1, 2$

1. La varianza de los rendimientos anuales de la acción i es un número que describe la "variabilidad" de estos rendimientos de un año a otro.

2. La covarianza de los rendimientos anuales de las acciones 1 y 2 es un número que describe o mide como los rendimientos de las dos acciones ascienden juntos.

3. El rendimiento esperado de la cartera se define por el número

$$x_1R_1 + x_2R_2.$$

4. La varianza de la cartera se define por el número

$$\sigma_1^2x_1^2 + 2\sigma_{12}x_1x_2 + \sigma_2^2x_2^2$$

El problema resulta ser un problema de programación cuadrática:

$$\text{mín}\, \sigma_1^2x_1^2 + 2\sigma_{12}x_1x_2 + \sigma_2^2x_2^2$$

$$\text{s. a.}$$

$$x_1R_1 + x_2R_2 \geq G$$

$$x_1 \leq S_1$$

$$x_2 \leq S_2$$

$$x_i \geq 0$$

4.2. Monopolista Discriminador

Una empresa vende un producto en dos áreas geográficas aisladas. Supongamos que puede poner precios diferentes en cada una de las áreas porque no es fácil revender en una lo que se ha comprado en otra. Supongamos también que la empresa tiene cierto poder de monopolio, en el sentido de ejercer una influencia sobre los precios en los dos mercados ajustando las cantidades en cada uno. Los economistas suelen llamar "monopolista discriminador" a una empresa con este poder.

Normalmente los precios en cada región son dadas por la fórmula siguiente (función inversa de la demanda)

$$P_1 = a_1 - b_1Q_1 \quad \text{y} \quad P_2 = a_2 - b_2Q_2$$

y los costos totales son proporcionales a la producción total, es decir

$$C(Q_1 + Q_2) = \alpha(Q_1 + Q_2) + \beta(Q_1 + Q_2)^2$$

Los beneficios totales estará dado por

$$\Pi(Q_1, Q_2) = P_1Q_1 + P_2Q_2 - \alpha Q_1 - \alpha Q_2 - \beta Q_1^2 - \beta Q_1^2 - \beta 2Q_1Q_2 - \beta Q_2^2.$$

El objetivo del monopolista es maximizar su beneficio, basándose en "Cuanto debe producir en cada región para obtener el máximo beneficio". Esto conduce a formular un problema cuadrático.

REFERENCIAS BIBLIOGRÁFICAS

[1] A. Friedlander - J.M. Martinez, *On the maximization of a concave quadratic function with box constrains*, SIAM Jornal on optimization 4, pp. 177-192 (1994).

[2] G.H. Golob - Ch. F. Van Loan, *Matrix computations*, The Johns Hopkins University Press, Baltimore and London (1989).

[3] I.I. Dikin , *Iterative Solution of Problems of Linear and Quadratic Programming*, Soviet Math, Dokl 8, pp. 674-675 (1967).

[4] J.F. Pierce - W.B. Crowston, *Tree - Search Algorithms for quadratic assigment problems*, Naval Renearch Logistics Quarterly, Volumen 18, Isswe 1, pp. 1-36, Date March (1971).

[5] J.V. Burke, *A sequential Quadratic Programming Method for Potentially Infeasible Mathematical Problems*, Journal of Mathematical Analysis and Applications 139, pp. 319-351 (1989).