



Juan Manuel Pacheco

Instituto de Investigaciones Teóricas y Aplicadas, Escuela de Estadística

PROGRAMACIÓN CUADRÁTICA Y SELECCIÓN DE CARTERAS DE INVERSIÓN

Introducción

El trabajo aquí presentado se origina en la tarea investigativa desarrollada por el autor en la cátedra "Matemáticas para Economistas II" de la carrera de Licenciatura en Economía, bajo la supervisión de la Lic. Nora Lac Prugent. Específicamente, este estudio se enmarca en la reorganización de dicha cátedra con la intención de mejorar la enseñanza de la matemática mediante la utilización de herramientas computacionales, a través del dictado de clases prácticas en el Laboratorio de Matemáticas de esta casa de estudios. Es por ello que el mismo forma parte del proyecto "La Ingeniería Didáctica en el Diseño y Seguimiento de Unidades Curriculares" dirigido por la Lic. Mercedes Anido.

Asimismo, una versión del presente trabajo obtuvo una mención especial en el concurso "Ing. Ricardo S. Carbajo" otorgado por la Asociación de Docentes de Matemática de Facultades de Ciencias Económicas y Afines en el año 2001.

Específicamente, lo que se plantea aquí es una aplicación de la enseñanza de la programación matemática, en especial la programación no lineal, mediante el desarrollo de un problema motivador vinculado al conocimiento afín del alumno. Esta modalidad de enseñanza permite articular los conocimientos teóricos – analíticos de un área de las matemáticas con los conceptos económicos adquiridos en otras materias, de tal forma que el alumno pueda profundizar en la interpretación, deducción y las variantes del problema en términos económico, en lugar de la exclusiva resolución matemática del mismo.

El desarrollo de los temas es el siguiente, en el primer apartado se presenta el problema motivador, tal como es presentado a los alumnos al inicio de la clase. A continuación se expone la teoría de programación no lineal, haciendo hincapié en las condiciones de Karush – Kuhn – Tucker. En el tercer apartado se presenta la teoría matemática de programación convexa y programación cuadrática, con la reinterpretación de los conceptos ya enunciados. Seguidamente es expuesta la teoría económica necesaria para la resolución del problema. En la siguiente sección se resuelve el problema motivador aplicando los conceptos teóricos mencionados previamente utilizando la herramienta Solver de Excel. En el sexto apartado, se presentan las consideraciones finales y conclusiones. Finalmente, las dos últimas secciones son el apéndice con conceptos y teoremas matemáticos y el apartado con la aplicación del Sistema DERIVE for Windows.



I. Problema motivador

Los objetivos del ejercicio son:

- 1) Lograr que los alumnos adquieran los conceptos matemáticos necesarios para resolver problemas de Programación No Lineal.
- 2) Lograr que los alumnos resuelvan problemas matemáticos mediante la Programación Matemática.
- 3) Lograr que los alumnos utilicen estos conocimientos para resolver problemas económicos.

A continuación se presenta el enunciado del problema motivador:

Un inversor adverso al riesgo desea invertir en un portfolio compuesto por las acciones A y B (siendo a y b las participaciones relativas de cada activo dentro de la cartera).

La matriz de variancias y covariancias asociada a estos activos es la siguiente:

$$V = \begin{pmatrix} 0,0081 & 0,02 \\ 0,02 & 0,0036 \end{pmatrix}$$

Asimismo, el inversor desea invertir todos los fondos y espera un rendimiento mínimo de la cartera del 3%, conociendo que los rendimientos esperados de los activos A y B son de 6% y 2%, respectivamente.

Se le solicita:

- a) Plantee el problema de programación no lineal con todas sus restricciones.
- b) Obtenga la composición de portfolio que le recomendaría al inversor.
- c) Grafique el problema planteado anteriormente.
- d) Estime si se cumplen las condiciones de Karush – Kuhn – Tucker.
- e) Determine si se satisface la cualificación de las restricciones.
- f) Determine si las condiciones Karush – Kuhn – Tucker son necesarias y/o suficientes.

II. Programación No Lineal con Restricciones

El problema general de programación no lineal restringido, puede definirse como:

Optimizar $z = f(X)$	siendo: f la función objetivo
Sujeto a: $g_i(X) \leq 0 \quad i: 1, 2, \dots, m$	g_i las restricciones de desigualdad
$g_i(X) \geq 0 \quad i: m+1, \dots, p$	h_i las restricciones de igualdad
$h_i(X) = 0 \quad i: p+1, \dots, r$	X el vector de variables de
$X \geq 0$	elección.

Se consideran que todas las funciones son dos veces diferenciable y al menos una de las mismas es no lineal.

No existe un algoritmo general para resolver modelos no lineales debido al comportamiento irregular de las funciones no lineales. Es por ello que en contraste con la programación lineal no se puede reducir el campo de elección al conjunto de puntos extremos de la región factible.



Sin embargo, se han determinado condiciones que bajo ciertos requisitos se convierten en condiciones de primer orden o necesarias, e inclusive en condiciones necesarias y suficientes. Estas son las condiciones de Karush – Kuhn – Tucker , que se indicarán como condiciones de KKT, y fueron desarrolladas independientemente por Karush y por Kuhn – Tucker.

Condiciones de Karush – Kuhn – Tucker

Dado que las condiciones KKT son el resultado analítico más importante en programación no lineal, se desarrollaran estas condiciones en dos pasos por conveniencia de exposición. En primer lugar se analizan las condiciones de no negatividad, para en el paso posterior desarrollar un problema con las condiciones de desigualdad tanto para maximización como para minimización.

a) Condiciones de no negatividad:

Como primer paso se considera un problema simple de optimización de la función $z = f(x_1)$ sujeta a la restricción que la variable de elección sea no negativa, es decir, $x_1 \geq 0$. Naturalmente esta condición es equivalente a $-x_1 \leq 0$, entonces incorporando una variable de holgura $s \geq 0$ y llamando μ al multiplicador de lagrange, la función lagrangiana resulta:

$$F(x_1, \mu, s) = f(x_1) + \mu (-x_1 + s) = 0$$

Las condiciones necesarias son:

$$\partial F / \partial x_1 = df/dx_1 - \mu = 0$$

$$\partial F / \partial \mu = -x_1 + s = 0$$

$$\partial F / \partial s = \mu s = 0$$

De la primera se desprende que $\mu = df/dx_1$. Entonces se pueden dar solamente tres tipos de extremos:

- Extremo local interno lo que implica que x_1 óptimo (x_1^*) es positivo y su derivada es nula, es decir que $\mu = 0$.
- Extremo local de frontera lo que implica que x_1^* es igual a cero y su derivada es nula
- Punto de frontera lo que implica que $x_1^* = 0$ y su derivada es positiva si es un problema de minimización, o negativa si es un problema de maximización.

Estas condiciones se pueden resumir de la siguiente manera:

$$df/dx_1 \leq 0 \quad x_1 \geq 0 \quad x_1 df/dx_1 = 0 \quad [\text{Maximización}]$$

$$df/dx_1 \geq 0 \quad x_1 \geq 0 \quad x_1 df/dx_1 = 0 \quad [\text{Minimización}]$$

Ampliando el problema para casos en que existen n variables de elección, es decir:

$$\text{Extremar} \quad z = f(x_1, x_2, \dots, x_n)$$

$$\text{sujeto a} \quad x_j \geq 0 \quad j: 1, 2, \dots, n$$

Las condiciones de primer orden para un óptimo son

$$df/dx_j \leq 0 \quad x_j \geq 0 \quad x_j df/dx_j = 0 \quad j: 1, 2, \dots, n \quad [\text{Maximización}]$$

$$df/dx_j \geq 0 \quad x_j \geq 0 \quad x_j df/dx_j = 0 \quad j: 1, 2, \dots, n \quad [\text{Minimización}]$$



b) Condiciones de desigualdad:

En este segundo paso se reconsidera el problema con la incorporación de una restricción de desigualdad y otra variable de elección. Además, dado que tradicionalmente los problemas de maximización se presentan con desigualdades del signo contrario a las de los problemas de minimización, se determinarán las condiciones para máximo y para mínimo a partir de problemas distintos.

Maximizar	$z = f(x_1, x_2)$	Minimizar	$z = f(x_1, x_2)$
sujeto a	$g_1(x_1, x_2) \leq r_1$	sujeto a	$g_1(x_1, x_2) \geq r_1$
	$x_j \geq 0 \quad j:1, 2$		$x_j \geq 0 \quad j:1, 2$

La restricción de desigualdad puede transformarse en una igualdad incorporando una variable de holgura o de excedente apropiada. Para este caso, la condición para máximo queda satisfecha sumando a $g(x_1, x_2) \leq r_1$ una variable de holgura no negativa s_1 ; asimismo para el problema de minimización se satisface restando a la función g una variable de excedente s no negativa.

De esta manera, las funciones de Lagrange sin considerar las restricciones de no negatividad serán:

$$F(x_1, x_2, \mu_1, s_1) = f(x_1, x_2) + \mu_1 [r_1 - g(x_1, x_2) - s_1] \quad [\text{Maximización}]$$

$$F(x_1, x_2, \mu_1, s_1) = f(x_1, x_2) + \mu_1 [r_1 - g(x_1, x_2) + s_1] \quad [\text{Minimización}]$$

y las condiciones necesarias de extremo se obtienen igualando a cero las derivadas parciales de F :

$$\partial F / \partial x_1 = (\partial f / \partial x_1) + \mu_1 (\partial g / \partial x_1) = 0$$

$$\partial F / \partial x_2 = (\partial f / \partial x_2) + \mu_1 (\partial g / \partial x_2) = 0$$

$$\partial F / \partial \mu_1 = 0$$

$$\partial F / \partial s_1 = 0$$

Y dado que las variables s_1 y x_j deben ser no negativas, de acuerdo a lo hallado en el punto a) relativo a las condiciones de no negatividad, se pueden reexpresar estas condiciones como:

$$\partial F / \partial x_j \leq 0 \quad x_j \geq 0 \quad x_j \partial F / \partial x_j = 0 \quad j: 1, 2$$

$$\partial F / \partial s_1 \leq 0 \quad s_1 \geq 0 \quad s \partial F / \partial s_1 = 0 \quad [\text{Maximización}]$$

$$\partial F / \partial \mu_1 = 0$$

$$\partial F / \partial x_j \geq 0 \quad x_j \geq 0 \quad x_j \partial F / \partial x_j = 0 \quad j: 1, 2$$

$$\partial F / \partial s_1 \geq 0 \quad s \geq 0 \quad s \partial F / \partial s_1 = 0 \quad [\text{Minimización}]$$

$$\partial F / \partial \mu_1 = 0$$

Luego, dado que $\partial F / \partial s_1 = -\mu_1$ para máximo y $\partial F / \partial s_1 = \mu_1$ para mínimo, la segunda línea determina que $-\mu_1 \leq 0$ para ambos tipos de problema, además s_1 es no negativo y por lo tanto dicha línea de condiciones es igual a

$$s_1 \geq 0 \quad \mu_1 \geq 0 \quad s_1 \mu_1 = 0$$



Pero a su vez la tercera línea implica que $s_1 = r_1 - g$ para máximo y $s_1 = -r_1 + g$. Por consiguiente, considerando la anterior se puede combinar la segunda con la tercera línea de tal forma que

$$r - g(x_1, x_2) \geq 0 \quad \mu_1 \geq 0 \quad \mu_1[r - g(x_1, x_2)] = 0 \quad [\text{Maximización}]$$

$$r - g(x_1, x_2) \leq 0 \quad \mu_1 \geq 0 \quad \mu_1[r - g(x_1, x_2)] = 0 \quad [\text{Minimización}]$$

Esto permite expresar las condiciones de KKT sin las variables artificiales.

$$\partial F / \partial x_j \leq 0 \quad x_j \geq 0 \quad x_j (\partial F / \partial x_j) = 0 \quad j: 1, 2$$

$$\partial F / \partial \mu_1 \geq 0 \quad \mu_1 \geq 0 \quad \mu_1 (\partial F / \partial \mu_1) = 0 \quad [\text{Maximización}]$$

$$\partial F / \partial x_j \geq 0 \quad x_j \geq 0 \quad x_j \partial F / \partial x_j = 0 \quad j: 1, 2$$

$$\partial F / \partial \mu_1 \leq 0 \quad \mu_1 \geq 0 \quad \mu_1 (\partial F / \partial \mu_1) = 0 \quad [\text{Minimización}]$$

Entonces para el caso general de n variables de elección y m restricciones, las funciones langragianas respectivas serán

$$F = f(x_1, x_2, \dots, x_n) + \sum_{i=1}^m \mu_i [r_i - g_i(x_1, x_2, \dots, x_n)] \quad [\text{Maximización}]$$

$$F = f(x_1, x_2, \dots, x_n) + \sum_{i=1}^m \mu_i [r_i - g_i(x_1, x_2, \dots, x_n)] \quad [\text{Minimización}]$$

Siendo las condiciones de KKT las siguientes

$$\partial F / \partial x_j \leq 0 \quad x_j \geq 0 \quad x_j (\partial F / \partial x_j) = 0$$

$$\partial F / \partial \mu_i \geq 0 \quad \mu_i \geq 0 \quad \mu_i (\partial F / \partial \mu_i) = 0 \quad [\text{Maximización}]$$

$$i: 1, 2, \dots, m \quad j: 1, 2, \dots, n$$

$$\partial F / \partial x_j \geq 0 \quad x_j \geq 0 \quad x_j \partial F / \partial x_j = 0$$

$$\partial F / \partial \mu_i \leq 0 \quad \mu_i \geq 0 \quad \mu_i (\partial F / \partial \mu_i) = 0 \quad [\text{Minimización}]$$

$$i: 1, 2, \dots, m \quad j: 1, 2, \dots, n$$

Verificándose cierta hipótesis sobre las restricciones, las condiciones de KKT son condiciones necesarias para un mínimo o máximo local, respectivamente. Y dado que todo extremo global debe ser un extremo local, estas condiciones son asimismo condiciones necesarias para extremos globales, siempre que se cumpla con la cualificación de las restricciones desarrolladas a continuación.

Cualificación de Restricciones

Existen ciertos requisitos sobre las restricciones de un programa no lineal para que las condiciones de KKT sean condiciones necesarias para un óptimo. Estos requisitos tienen la intención de salvar irregularidades que pueden existir en la frontera de la región factible cuando las restricciones son no lineales.



Considerando un punto de frontera que es candidato para una solución $X^* \equiv (x_1^*, x_2^*, \dots, x_n^*)$, y el vector de diferenciales $dX \equiv (dx_1, dx_2, \dots, dx_n)$ que indica el movimiento en una dirección específica a partir del punto de frontera X^* .

La cualificación de las restricciones establece dos condiciones para los vectores dX :

- Si la j -ésima variable de elección tiene valor cero en el punto X^* , entonces $dx_j \geq 0$.
- Si la i -ésima restricción se satisface como igualdad en el punto X^* , entonces $dg_i(X^*) = (\partial g/\partial x_1) dx_1 + (\partial g/\partial x_2) dx_2 + \dots + (\partial g/\partial x_n) dx_n \leq 0$ para problemas de maximización y $dg_i(X^*) \geq 0$ para minimización; donde todas las derivadas parciales se calculan en X^* .

Todo vector que satisfaga a) y b) es considerado un vector prueba. Finalmente si existe un arco diferenciable que procede del punto X^* , está enteramente contenido en la región factible y es tangente al vector prueba dado, se le denomina arco de cualificación para dicho vector.

Con ello la cualificación de las restricciones se satisface si para cualquier punto X^* sobre la frontera de la región factible existe un arco de cualificación para cada vector de prueba.

III. Programación Convexa y Programación Cuadrática

Un programa de optimización convexo, resuelve dos problemas similares:

- Minimizar una función convexa cuando el espacio de soluciones es un conjunto convexo.
- Maximizar una función cóncava siendo la región factible un conjunto convexo.¹

La importancia del análisis de este tipo de programación obedece a posibilidad de transformar las condiciones de KKT en condiciones suficientes para óptimos. Para ello es necesario enunciar los siguientes teoremas sobre óptimos locales y globales.

Teorema 1: Sean dados el conjunto convexo S de \mathfrak{R}^n y la función convexa $f(X)$ en S . Dado el problema de minimizar $f(X)$ sujeto a $X \in S$, si X_0 es una solución óptima local, entonces X_0 es una solución óptima global. Además, si $f(X)$ es estrictamente convexa, X_0 es la única solución óptima

Demostración: Puesto que X_0 es una solución óptima local, entonces existe un entorno $N(X_0)$ tal que

$$f(X) \geq f(X_0) \quad \forall X \in N(X_0) \cap S \quad (i)$$

Suponiendo que X_0 no es una solución óptima global, de modo que existe $X_1 \in S$ tal que $f(X_1) < f(X_0)$. Pero por la convexidad de f , para un $\alpha \in (0, 1)$ se cumple que:

$$f[\alpha X_1 + (1 - \alpha)X_0] \leq \alpha f(X_1) + (1 - \alpha) f(X_0) < \alpha f(X_0) + (1 - \alpha) f(X_0) = f(X_0) \quad (ii)$$

Ahora, para $\alpha > 0$ suficiente pequeño

$$X = \alpha X_1 + (1 - \alpha)X_0 \in N(X_0) \cap S$$

y en consecuencia la desigualdad (ii) contradice la (i) y la primera parte del teorema queda demostrada por reducción al absurdo.

¹ Ver definiciones en las secciones B y C del apéndice.



Finalmente, suponiendo que X_0 es una solución óptima local y $f(X)$ es estrictamente convexa, lo que implica que $f(X)$ es convexa, se deduce por lo demostrado anteriormente que X_0 es una solución global.

Contrariamente, suponiendo que X_0 no es la única solución óptima global, entonces existe $X_2 \in S$ y $X_2 \neq X_0$ tal que $f(X_2) = f(X_0)$. Pero por la estricta convexidad de f será:

$$f(\frac{1}{2} X_2 + \frac{1}{2} X_0) < \frac{1}{2} f(X_2) + \frac{1}{2} f(X_0) = f(X_0)$$

Ahora, por la convexidad de S , $(\frac{1}{2} X_2 + \frac{1}{2} X_0) \in S$ y la desigualdad anterior viola la optimalidad global de X_0 .

En consecuencia X_0 es el único punto en que f alcanza el mínimo global o absoluto.

Nota: El caso de maximizar una función cóncava es similar al de minimizar una función convexa. En consecuencia, para el problema de maximizar una función cóncava $f(X)$ sujeta a $X \in S$, siendo S un conjunto convexo de \mathfrak{R}^n , existe un teorema análogo al precedente.

Teorema 2: Dados $f: \mathfrak{R}^n \rightarrow \mathfrak{R}$ una función convexa diferenciable en \mathfrak{R}^n y S un conjunto convexo de \mathfrak{R}^n , se considera el problema de minimizar $f(X)$ sujeto a $X \in S$. Entonces, $X_0 \in S$ es una solución óptima si y solo si $[\nabla f(X_0)]^T (X - X_0) \geq 0 \quad \forall X \in S$. Además, si S es un conjunto abierto, X_0 es una solución óptima si y solo si $\nabla f(X_0) = 0$.

Demostración: Si $[\nabla f(X_0)]^T (X - X_0) \geq 0 \quad \forall X \in S$, como por convexidad de f se tiene

$$f(X) \geq f(X_0) + [\nabla f(X_0)]^T (X - X_0) \quad \forall X \in S,$$

resulta entonces $f(X) \geq f(X_0) \quad \forall X \in S$ (iii); por tanto X_0 es una solución óptima del problema.

Recíprocamente, sea $X_0 \in S$ una solución óptima. Por la convexidad de S , $\forall X \in S$ y $\alpha \in [0, 1]$ tenemos que $\alpha X + (1 - \alpha)X_0 \in S$, o sea $X_0 + \alpha(X - X_0) \in S$ y la desigualdad (iii) puede reescribirse

$$f[X_0 + \alpha(X - X_0)] - f(X_0) \geq 0 \quad \forall X \in S \quad \text{y} \quad 0 \leq \alpha \leq 1,$$

y siendo f diferenciable se tiene:

$$f[X_0 + \alpha(X - X_0)] - f(X_0) = \alpha[\nabla f(X_0)]^T (X - X_0) + \alpha \bullet X - X_0 \bullet \bullet (\alpha (X - X_0))$$

luego: $\alpha [\nabla f(X_0)]^T (X - X_0) + \alpha \bullet X - X_0 \bullet \bullet (\alpha (X - X_0)) \geq 0$

Dividiendo por α resulta:

$$[\nabla f(X_0)]^T (X - X_0) + \bullet X - X_0 \bullet \bullet (\alpha (X - X_0)) \geq 0$$

y tomando límite para $\alpha \rightarrow 0$

$$[\nabla f(X_0)]^T (X - X_0) \geq 0$$

Finalmente, si S es abierto todo punto es interior al mismo, luego si en X_0 la función alcanza un mínimo, por las conocidas condiciones necesarias de extremo, las derivadas parciales, que existen por la diferenciabilidad de f en S , deberán anularse en X_0 , luego $\nabla f(X_0) = 0$ (iv).

Recíprocamente, si se verifica (iv), X_0 es una solución óptima. En efecto, de la desigualdad

$$f(X) \geq f(X_0) + [\nabla f(X_0)]^T (X - X_0) \quad \forall X \in S$$

se desprende que

$$f(X) \geq f(X_0) \quad \forall X \in S$$

completándose así la demostración.

Suficiencia de las condiciones de KKT

Las condiciones de KKT son suficientes en un problema de programación convexo.



Dado que es más sencillo verificar que una función sea convexa o cóncava que demostrar que una región de factibilidad es un conjunto convexo, se consideran las siguientes definiciones y corolarios surgidos de los teoremas enunciados en las secciones A y B del apéndice:

Definición 1: Siendo $A = (a_1, a_2, \dots, a_n)$, $X = (x_1, x_2, \dots, x_n)^T$ y b un número real, se llama hiperplano en \mathfrak{R}^n al conjunto de puntos que satisfacen la ecuación lineal $AX = b$. Todo hiperplano divide los puntos \mathfrak{R}^n en dos semiespacios: $AX \leq b$ y $AX \geq b$

Corolario 1: Todo semiespacio de \mathfrak{R}^n es un conjunto convexo.

Corolario 2: El hiperplano $AX = b$ es un conjunto convexo.

Corolario 3: Si la función $g(X)$ es convexa en S convexo, entonces el conjunto $T = \{X \in S : g(X) \leq k\}$ es convexo.

Corolario 4: Si la función $g(X)$ es cóncava en S convexo, entonces el conjunto $T = \{X \in S : g(X) \geq k\}$ es convexo.

Corolario 5: Si $h(X)$ es una función lineal definida en S convexo, entonces $h(X)$ es a la vez convexa y cóncava.

De esta forma la convexidad del espacio de soluciones se puede establecer verificando directamente la convexidad o la concavidad de las funciones que constituyen las restricciones.

Para el caso general de Maximización:

$$\begin{array}{ll} \text{Maximizar} & z = f(x_1, x_2, \dots, x_n) \\ \text{sujeto a} & g_i(x_1, x_2, \dots, x_n) \leq r_i \quad i: 1, 2, \dots, m \\ & x_j \geq 0 \quad j: 1, 2, \dots, n \end{array}$$

El problema será de programación convexa si y solo si:

- La función objetivo $f(x_1, x_2, \dots, x_n)$ es cóncava en el ortante no negativo.
- Las funciones $g_i(x_1, x_2, \dots, x_n)$ son convexas en el ortante no negativo.

Para el caso general de Minimización:

$$\begin{array}{ll} \text{Minimizar} & z = f(x_1, x_2, \dots, x_n) \\ \text{sujeto a} & g_i(x_1, x_2, \dots, x_n) \geq r_i \quad i: 1, 2, \dots, m \\ & x_j \geq 0 \quad j: 1, 2, \dots, n \end{array}$$

El problema será de programación convexa si y solo si:

- La función objetivo $f(x_1, x_2, \dots, x_n)$ es convexa en el ortante no negativo.
- Las funciones $g_i(x_1, x_2, \dots, x_n)$ son cóncavas en el ortante no negativo.

Las condiciones de KKT como condiciones necesarias y suficientes para extremos

De lo anterior se puede deducir que las condiciones de KKT son condiciones necesarias y suficientes si:

- El problema es de programación convexa.
- El punto óptimo cumple con la cualificación de las restricciones.



Programación Cuadrática

La programación cuadrática considera el problema de optimizar una función objetivo cuadrática sujeta a restricciones lineales y a condiciones de no negatividad. Este tipo de programación suele ser muy importante en el estudio de las Ciencias Económicas dado que las formulaciones de programas cuadráticos surgen de manera natural en muchas aplicaciones.

Un modelo de programación cuadrática se define de la siguiente manera:

$$\text{Optimizar } f(X) = CX + X^TDX$$

$$\text{Sujeto a: } AX \leq B$$

$$X \geq 0$$

$$\text{siendo: } X = (x_1, x_2, \dots, x_n)^T$$

$$C = (c_1, c_2, \dots, c_n)$$

$$B = (b_1, b_2, \dots, b_m)^T$$

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{r1} & a_{r2} & \dots & a_{rn} \end{pmatrix}$$

$$D = \begin{pmatrix} d_{11} & d_{12} & \dots & d_{1n} \\ d_{21} & d_{22} & \dots & d_{2n} \\ \dots & \dots & \dots & \dots \\ d_{n1} & d_{n2} & \dots & d_{nn} \end{pmatrix}$$

Al cual se le pueden establecer ciertos requisitos para que las condiciones de KKT sean condiciones necesarias y suficientes.

Primero, se debe destacar que la linealidad de las restricciones garantiza que el espacio de soluciones sea un conjunto convexo.

Luego el problema queda reducido a determinar la concavidad o convexidad estricta de la función objetivo f de acuerdo si el problema es de maximización o minimización, respectivamente.

A partir de la definición de gradiente dada en la sección A-3) del apéndice se tiene que:

$$\nabla f(X) = \nabla(CX + X^TDX) = C + 2DX$$

Siendo la matriz hessiana H de $f(X)$:

$$H = \nabla^2 f(X) = \nabla(C + 2DX) = 2D$$

Con lo cual si la matriz D es definida positiva, H también lo será, en cuyo caso f resultará estrictamente convexa (ver teorema a) de la sección C-3) del apéndice).

De la anterior se deduce que en un problema de minimización las condiciones KKT son condiciones necesarias y suficientes si y solo si la matriz D es definida positiva, lo que equivale a que los autovalores de dicha matriz sean todos positivos de acuerdo al teorema c) de la sección C-3) del apéndice.

Asimismo, si el problema es de maximización las condiciones de KKT serán condiciones necesarias y suficientes si y solo si la matriz D es definida negativa. Lo que es equivalente a que la $f(X)$ es estrictamente cóncava, siendo los autovalores de la matriz D todos negativos.

Por las condiciones expuestas, estos problemas son problemas de programación convexa, aplicándose a ello todo lo expuesto en las secciones previas.



IV. Selección de Carteras de Activos Financieros

Introducción

Los conceptos que se presentan a continuación constituyen la base de la teoría moderna de las finanzas corporativas. Esencialmente, este apartado intenta deducir la relación entre las dos características básicas de un instrumento financiero: la rentabilidad y el riesgo.

La rentabilidad de un título se la puede definir como el beneficio económico que se obtiene a partir de la tenencia de dicho título durante un período determinado. Generalmente, se expresa en porcentajes. A partir de la definición se deduce que la rentabilidad de un activo financiero es solo conocida luego que el período en el que se esté evaluando haya vencido, es decir, que no se puede fijar con exactitud cual ha de ser la rentabilidad del mismo en el futuro. Esta deducción puede asimismo considerarse en virtud de la relación inversa existente entre el precio de un activo y su rentabilidad. Con lo cual, dado que el precio exacto de los títulos, al igual que el de cualquier otra mercancía, es incierto hacia el futuro, entonces tampoco puede ser conocida su rentabilidad.

Es por ello que se suelen considerar dos términos distintos aunque muy relacionados: rentabilidad histórica y rentabilidad esperada. La rentabilidad histórica de un activo financiero es la rentabilidad que el mismo ha reportado a lo largo de su existencia; generalmente se presentan las tasas de rentabilidad histórica por año y se construye un histograma con las mismas. Esta información suele ser la base para estimar las tasas de rentabilidad esperada del mismo título para los períodos futuros, mediante diferentes tipos de modelos.

Asimismo, la incertidumbre da origen a la otra característica importante de cualquier activo financiero: su riesgo. El riesgo es la medida de la posibilidad de perder o no poder ganar cierto valor económico. Esta característica obedece a que los activos financieros son títulos contingentes, es decir, compromisos de pagos futuros que no son 100% seguros.

A pesar de ser clara su definición, el riesgo es una medida de difícil interpretación en la práctica. Una manera de considerar el riesgo de la rentabilidad de los instrumentos financieros es en términos de la dispersión de la distribución de frecuencia de las tasas de rentabilidad históricas. La dispersión de la distribución es una medida de cuánto se puede desviar una rentabilidad determinada de la rentabilidad media. Es por ello que se utilizan dos conceptos estadísticos como medidas de riesgo de un título: la varianza (σ^2 o Var) y su raíz cuadrada, el desvío estándar (SD o σ).

Carteras y Diversificación

Las carteras de títulos o portfolios son combinaciones de activos financieros que se realizan con la intención de reducir el riesgo que estos poseen individualmente. A este fenómeno se lo conoce como diversificación y su eficiencia para reducir el riesgo se evidencia en las grandes diferencias que existen entre la desviación estándar de un título individual y la desviación estándar de una cartera o un índice.²

² Por ejemplo, la desviación estándar histórica del índice compuesto S&P 500 (que abarca 500 de los capitales sociales más grandes de EE.UU.) es del 20,8% mientras que la SD de la compañía Chrysler es 47% y el de Apple es del 38%.



Con la opción de diversificar, todo inversor adverso al riesgo invertirá en carteras antes que en activos individuales, dada la posibilidad de obtener las mismas rentabilidades pero con menor riesgo. Es por ello que el análisis de los títulos se hace desde el punto de vista de la contribución de cada título a la rentabilidad esperada y al riesgo de la cartera.

Dado que la rentabilidad esperada de la cartera es el promedio aritmético de las rentabilidades esperadas de los activos que la componen, la rentabilidad esperada de un título es la medida adecuada de la contribución del título a la rentabilidad esperada de la cartera. Sin embargo, ni la varianza ni la desviación estándar son medidas adecuadas de la contribución de un título al riesgo de una cartera cuando la correlación entre cada par de activos que componen la cartera no es perfecta.

Para poder entenderlo con mayor claridad conviene analizar la fórmula de la varianza de una cartera compuesta por 2 títulos, la misma es:

$$\text{Var (cartera)} = x_1^2 \sigma_1^2 + 2x_1 x_2 \sigma_{1,2} + x_2^2 \sigma_2^2 = x_1^2 \sigma_1^2 + 2x_1 x_2 \rho_{1,2} \sigma_1 \sigma_2 + x_2^2 \sigma_2^2,$$

donde x_i es la participación del título i en el portfolio, σ_i^2 es la varianza del título i , y $\rho_{i,j}$ es la correlación entre los activos i y j .

Al observar la expresión de la derecha se puede deducir que únicamente cuando $\rho_{i,j}$ es igual a 1 (correlación perfecta) la SD de la cartera es un promedio ponderado de las desviaciones estándar de las rentabilidades.; en el caso en que no exista correlación perfecta, el desvío estándar del portfolio será menor al promedio de los SD de los títulos, es decir, tendrá menor variabilidad y por ende menor riesgo.

Puede asimismo hallarse la medida de riesgo para una cartera con n títulos (por simplificación se utiliza notación matricial):

$$\sigma_{\text{cart}}^2 = \mathbf{X}^T \mathbf{V} \mathbf{X},$$

siendo: $\mathbf{X} = (x_1, x_2, \dots, x_n)^T$ el vector de participación relativa de cada título en la cartera,

$$\mathbf{V} = \begin{pmatrix} \sigma_1^2 & \sigma_{1,2} & \sigma_{1,3} & \dots & \sigma_{1,n} \\ \sigma_{2,1} & \sigma_2^2 & \sigma_{2,3} & \dots & \sigma_{2,n} \\ \sigma_{3,1} & \sigma_{3,2} & \sigma_3^2 & \dots & \sigma_{3,n} \\ \dots & \dots & \dots & \dots & \dots \\ \sigma_{n,1} & \sigma_{n,2} & \sigma_{n,3} & \dots & \sigma_n^2 \end{pmatrix} \text{ la matriz de varianzas y covariancias.}$$

Es evidente que mientras mayor sea la cantidad de activos en la cartera, mayor es la influencia sobre σ_{cart}^2 de las covariancias entre activos y menor la de las variancias de cada uno. Esto demuestra que en carteras bien diversificadas lo que adiciona cada título al riesgo de la cartera depende en mayor medida de la forma en que varía su rentabilidad respecto a la de los otros activos que conforman la cartera y en menor medida de la variación de la rentabilidad de ese título con su propia rentabilidad promedio.

Asimismo, como $\sigma_{ij} = \sigma_{ji}$, entonces la matriz \mathbf{V} es simétrica y σ_{cart}^2 es una forma cuadrática. Esto permite establecer problemas de programación matemática, específicamente de programación cuadrática para resolver el problema de agentes adversos al riesgo que consideran minimizar su riesgo (representado por la variancia de la cartera) sujeto a obtener un mínimo de rentabilidad.



Por lo tanto, si se considera que $E(R_i)$ es la rentabilidad esperada del título i y R_{cart} es la rentabilidad mínima que se desea obtener de la cartera, entonces continuando con la notación anterior, se puede establecer el problema de selección de cartera de mínima variancia de acuerdo al siguiente programa cuadrático:

$$\begin{aligned} \text{Minimizar} \quad & \sigma_{cart}^2 = X^T V X \\ \text{sujeto a:} \quad & E(R_1) + E(R_2) + \dots + E(R_n) \leq R_{cart} \\ & x_1 + x_2 + \dots + x_n = 1 \\ & x_1, x_2, \dots, x_n \geq 0, \end{aligned}$$

donde la primera restricción obedece a la rentabilidad mínima deseada y la siguiente al requisito de inversión total de fondos.

Para poder resolver este tipo de problemas, especialmente si la cartera contiene una gran cantidad de activos, conviene utilizar algún software como el Excel de Microsoft o LINDO.

A continuación se muestra un ejemplo de resolución utilizando la herramienta solver de Excel, cabe recordar que la solución que sigue corresponde al enunciado presentado en la sección II.

V. Resolución del Problema Motivador

a) El planteamiento del problema es el siguiente:

$$\begin{aligned} \text{Minimizar} \quad & \sigma(a,b) = 0,0081a^2 + 0,04ab + 0,0036b^2 \\ \text{sujeto a:} \quad & 0,06 a + 0,02 b \geq 0,03 \\ & a + b = 1 \\ & a, b \geq 0 \end{aligned}$$

b) La solución del problema fue hallada utilizando la herramienta de Solver de la planilla de cálculo de Excel de acuerdo a los siguientes pasos:

1- Se utilizan dos columnas: la A y la B. En la primera de ella se establece el planteamiento del problema tal cual lo ve el alumno. La B es utilizada para establecer las formulas con las que trabaja Solver, para ello hay que utilizar fórmulas para el planteamiento de la función objetivo y las restricciones. Estas van a estar en función de las celdas que representan las variables; en nuestro ejemplo son las celdas B7 y B8.

2- Paso seguido se selecciona Solver... del menú de Herramientas lo que despliega un cuadro de diálogo.

3- En dicho cuadro de diálogo se deben seleccionar la celda objetivo (en nuestro caso es B4); si se desea maximizar, minimizar o dar un valor a la celda objetivo; y las celdas que contienen las variables (B7 y B8); y finalmente se introducen las restricciones. Como Solver tiene una opción para las condiciones de no negatividad en el paso anterior estas no son introducidas.

4- Posteriormente, se debe realiza un clic con el mouse sobre Opciones lo cual despliega un nuevo cuadro de diálogo. Este contiene diversas opciones como tiempo máximo de cálculo, precisión, iteraciones, etc.



5- En este cuadro lo relevante se ubica en la parte inferior izquierda. En primer lugar se debe desactivar (si es que está con un tilde) la opción de "Adoptar modelo lineal". Luego se debe activar con un tilde la opción "Asumir no negativos" que equivale a introducir las condiciones de no negatividad de las variables de elección. Finalmente en la sección titulada Estimación se debe seleccionar Cuadrática en lugar de Lineal. Establecido todo lo anterior se da aceptar.

6- Seguidamente, al volver al cuadro de Solver, se debe seleccionar Resolver.

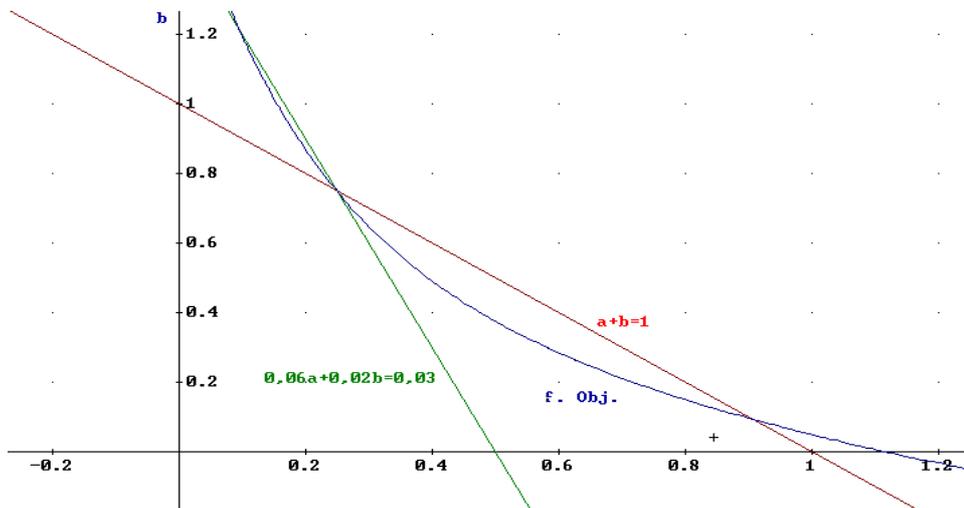
7- Finalmente Excel informa si se ha alcanzado una solución y da la opción a utilizar dicha solución hallada o volver a los valores iniciales de las celdas. Asimismo, da la posibilidad de obtener 3 informes: Respuestas, Sensibilidad y Límites.

En el ejercicio que se está resolviendo, la solución dada por Excel es la siguiente:

Problema Motivador	
Función Objetivo	
$r = (0,09a)^2 + 0,04ab + (0,06b)^2$ 0,01003126	
Variables	
A	0,2499995
B	0,7500015
Restricciones	
A + b = 1	1,000001
$0,06a + 0,02b \geq 0,03$	0,03

Se debe resaltar que dado la opción a Excel de un grado de error de acuerdo a la precisión fijada, alguna de las restricciones puede no darse estrictamente, tal es el caso de la restricción $a + b = 1$ en este ejercicio.

c) Para realizar el gráfico se utilizó DERIVE for Windows. El mismo se muestra a continuación y presenta tanto las restricciones como la función objetivo valorada en el punto mínimo.





- d) Para determinar si se cumplen las condiciones de KKT se debe armar la función lagrangiana, pero antes de ello se debe reexpresar el problema de tal manera que las restricciones sean todas mayores o iguales dado que es un problema de minimización, es decir:

$$\begin{aligned} \text{Minimizar } \quad & \sigma(a,b) = 0,0081a^2 + 0,04ab + 0,0036b^2 \\ \text{sujeto a: } \quad & 0,06 a + 0,02 b \geq 0,03 \\ & a + b \geq 1 \\ & - a - b \geq -1 \\ & a, b \geq 0 \end{aligned}$$

De esta forma el lagrangiano es:

$$L = 0,0081 a^2 + 0,04 a b + 0,0036 b^2 + \mu_1 (1 - a - b) + \mu_2 (-1 + a + b) + \mu_3 (0,03 - 0,06 a - 0,02 b)$$

Derivando respecto a las variables de elección (a y b) y los multiplicadores de lagrange, y considerando el valor óptimo de las variables, se tiene:

$$\partial F / \partial a = 0,03405 - \mu_1 + \mu_2 - 0,06 \mu_3$$

$$\partial F / \partial b = 0,0154 - \mu_1 + \mu_2 - 0,02 \mu_3$$

$$\partial F / \partial \mu_1 = 1 - a - b$$

$$\partial F / \partial \mu_2 = -1 + a + b$$

$$\partial F / \partial \mu_3 = 0,03 - 0,06 a - 0,02 b$$

Como ambas variables de elección son distintas de cero en el óptimo, entonces las dos primeras derivadas deben ser iguales a cero, con lo cual se obtiene por despejar en ambas ecuaciones el valor de μ_3 y la relación entre μ_1 y μ_2 :

$$\mu_3 = 0,046625$$

$$\mu_1 = 0,006075 + \mu_2$$

Como todas las restricciones se cumplen como igualdad, entonces todas las derivadas con respecto a los multiplicadores de lagrange son iguales a cero, esto indica que los μ_i deben ser todos no negativos para que se cumplan las condiciones de KKT. En este caso el valor de μ_3 es positivo y dando cualquier valor positivo a μ_2 se obtiene μ_1 positivo. En conclusión, se satisfacen las condiciones de Karush – Kuhn – Tucker.

- e) Para el análisis de la cualificación de las restricciones, se debe tener en cuenta cual es la región factible. En este caso, la misma comprende el segmento de la recta de la restricción $a + b = 1$ desde el punto óptimo ($a^* = 0,25$; $b^* = 0,75$) hasta el punto en que dicha recta corta el eje de las abscisas ($a = 1$; $b = 0$).

A partir de ello se deben hallar las restricciones sobre los diferenciales de las variables de elección para construir los vectores pruebas.

En primer lugar, dado que ninguna de las variables es igual a cero en el óptimo, no se puede imponer que sus diferenciales sean no negativos.



En cambio, como todas las restricciones se cumplen como igualdad, se deben realizar las siguientes tres diferenciaciones (recordar que dos de estas tres restricciones corresponden a una sola de las restricciones originales, que ha tenido que ser reexpresada para poder realizar el punto anterior y el presente):

$$\begin{aligned} da + db &\geq 0 \\ -da - db &\geq 0 \\ 0,06 da + 0,02 db &\geq 0 \end{aligned}$$

De las dos primeras se deduce que $db = -da$; de la tercera se obtiene que $db \geq -3 da$.

A partir de ello, se puede comenzar a armar los vectores pruebas. Por ejemplo el vector $(da; db) = (1; -1)$ si bien cumple con la primera relación hallada, no cumple con la segunda, con lo cual no es un vector prueba. De hecho ningún vector cuya componente db sea negativa es un vector prueba.

Sin embargo, el vector $(-0,05; 0,05)$ es un vector prueba que posee un arco de cualificación. Asimismo, se pueden encontrar infinidad de vectores pruebas, poseyendo todos ellos un arco de cualificación. Esto podría deducirse igualmente al considerar que todas las restricciones son lineales, con lo cual siempre se satisfacen la cualificación de las restricciones.

- f) A partir de lo hallado en el punto anterior se puede afirmar que las condiciones de KKT son condiciones necesarias.

Para conocer si dichas condiciones son además condiciones suficientes se debe determinar si el problema es un programa convexo. Para ello hay que verificar que tanto el espacio de soluciones como la función objetivo sean convexos, dado que es un problema de minimización.

Siendo todas las restricciones lineales se puede afirmar que la región factible es un espacio convexo.

Asimismo, para corroborar si la función objetivo es convexa se puede analizar los menores principales o los autovalores de la matriz de variancias y covariancias. Los resultados expuestos a continuación se hallaron utilizando el programa DERIVE for Windows³:

$$\begin{aligned} \text{Menores principales:} \quad v_{11} &= 0,0081 > 0 \\ |V| &= -0,00037 > 0 \\ \text{Autovalores } (\lambda): \quad \lambda_1 &= 0,0259761 > 0 \\ \lambda_2 &= -0,142761 < 0 \end{aligned}$$

Con lo cual se deduce que no es convexa la función objetivo dado que tanto los autovalores como los menores principales muestran que la matriz V es indefinida; es decir, no es definida positiva como requiere la teoría.

³ La aplicación de DERIVE for Windows que ha sido utilizada para extraer esta información se encuentra en un apartado al final.



VI. Consideraciones finales y conclusiones

El presente trabajo surge como una propuesta alternativa a los métodos tradicionales de enseñanza de la matemática, especialmente apropiado para carreras que utilizan la matemática como herramienta. La intención del mismo es presentar uno de los innumerables casos en que la teoría matemática sirve como instrumento para la resolución de problemas de economía aplicada.

Con tal propósito, se ha intentado exhibir una metodología apropiada para las materias de Matemáticas en las carreras de Ciencias Económicas (Economía, Administración de Empresas, Contaduría, etc.), la cual hace hincapié en la profunda interrelación entre la matemática y las teorías económicas que se imparten en dichas carreras.

Finalmente, se debe destacar que por la experiencia en el aula se puede concluir que los alumnos que han participado activamente en el proceso de aprendizaje, relacionando los conceptos matemáticos con los económicos y utilizando un software que permita desviar la atención de la resolución hacia el análisis del problema, han presentado un mayor interés en los temas matemáticos y han profundizado los mismos con la finalidad de resolver problemas económicos más complejos relacionados con el problema motivador propuesto inicialmente.

VII. Apéndice

A- Elementos de Algebra y Análisis

1) Formas Cuadráticas

a) Definición: dados

$$X = (x_1, x_2, \dots, x_n)^T$$

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

La función real definida en \mathfrak{R}^n :

$$Q(X) = X^T A X$$

se denomina forma cuadrática y es un polinomio homogéneo de segundo grado

Siendo evidente que toda forma cuadrática se anula en el origen, es decir, si $X = 0 \therefore Q(0) = 0$.

Interesa por lo tanto el comportamiento de Q fuera del origen.

De tal forma se puede establecer que la forma cuadrática es:

- Definida positiva si $Q(X) > 0 \forall X \neq 0$
- Definida negativa si $Q(X) < 0 \forall X \neq 0$
- Semidefinida positiva si $Q(X) \geq 0 \forall X \in \mathfrak{R}^n$ (en este caso $\exists X \neq 0 / Q(X) = 0$)
- Semidefinida negativa si $Q(X) \leq 0 \forall X \in \mathfrak{R}^n$
- Indefinida, cuando Q toma valores positivos y valores negativos en \mathfrak{R}^n

b) Teorema de Sylvester: Para que una forma cuadrática Q(X) sea definida positiva es necesario y suficiente que sean positivos todos los menores principales del determinante de A, es decir:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{11} & a_{12} & a_{13} \end{vmatrix}$$



$$\Delta_1 = a_{11} > 0; \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0; \Delta_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} > 0; \dots; \Delta_n |\mathbf{A}| > 0$$

c) Condición necesaria y suficiente para que la forma cuadrática $Q(X)$ sea definida negativa, es que $-Q(X)$ sea definida positiva. Claro que para que $-Q(X)$ sea definida positiva, según el criterio de Sylvester deberán ser:

$$\Delta_1 = -a_{11} > 0; \Delta_2 = \begin{vmatrix} -a_{11} & -a_{12} \\ -a_{21} & -a_{22} \end{vmatrix} > 0; \Delta_3 = \begin{vmatrix} -a_{11} & -a_{12} & -a_{13} \\ -a_{21} & -a_{22} & -a_{23} \\ -a_{31} & -a_{32} & -a_{33} \end{vmatrix} > 0; \dots; \Delta_n |-\mathbf{A}| > 0$$

pero esto significa que:

$$\Delta_1 = a_{11} < 0; \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0; \Delta_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} < 0; \dots; (-1)^n \Delta_n > 0$$

2) Matrices definidas y semidefinidas

Sea A una matriz simétrica $n \times n$. Se dice que A es definida positiva si $X^T A X > 0 \forall X \neq 0$ de \mathfrak{R}^n .

Asimismo, A es semidefinida positiva si $X^T A X \geq 0 \forall X \in \mathfrak{R}^n$.

Si $X^T A X < 0 \forall X \neq 0$ entonces se dice que A es definida negativa y si $X^T A X \leq 0 \forall X \in \mathfrak{R}^n \therefore A$ es semidefinida negativa.

3) Funciones Diferenciables

Indicando con S un conjunto abierto de \mathfrak{R}^n . Sea f una función definida en S , se define a f como diferenciable en $X_0 \in S$ si existe un vector $\nabla f(X_0)$ en \mathfrak{R}^n llamado gradiente de f en X_0 y una función $\omega: \mathfrak{R}^n \rightarrow \mathfrak{R}$ tal que:

$$f(X) = f(X_0) + [\nabla f(X_0)]^T (X - X_0) + \bullet(X - X_0) \bullet \bullet(X_0; (X - X_0)) \forall X \in S$$

donde $\lim_{\bullet(X; X - X_0) = 0} \bullet(X; X - X_0) = 0$ cuando X tiende a X_0 .

La función f es diferenciable en S , si lo es en cada punto de S .

Asimismo, se define vector gradiente al vector cuyos elementos son las derivadas parciales primeras, es decir:

$$\nabla f(X_0) = [(\partial f(x_0)/\partial x_1); (\partial f(x_0)/\partial x_2); \dots; (\partial f(x_0)/\partial x_n)]^T$$

4) Funciones dos veces diferenciables

Sean S un conjunto abierto en \mathfrak{R}^n y $f: S \rightarrow \mathfrak{R}$. Se dice que f es dos veces diferenciables en $X_0 \in S$, si además del vector gradiente, existe una matriz $n \times n$ simétrica $H(X_0)$ llamada matriz hessiana de f en X_0 y un función $\omega: \mathfrak{R}^n \rightarrow \mathfrak{R}$, tal que:

$$f(X) = f(X_0) + [\nabla f(X_0)]^T (X - X_0) + \frac{1}{2} (X - X_0)^T H(X_0) (X - X_0) + \bullet(X - X_0)^2 \bullet(X_0; X - X_0) \forall X \in S$$

donde $\lim_{\bullet(X_0; X - X_0) = 0} \bullet(X_0; X - X_0) = 0$ cuando $X \rightarrow X_0$.

La matriz hessiana formada por las derivadas parciales segundas es.

$$H(X_0) = \begin{pmatrix} (\partial^2 f(X_0)/\partial X_1^2) & (\partial^2 f(X_0)/\partial X_1 \partial X_2) \dots & (\partial^2 f(X_0)/\partial X_1 \partial X_n) \\ \dots & \dots & \dots \\ \dots & \dots & \dots \end{pmatrix}$$



$$(\partial^2 f(X_0)/\partial X_n \partial X_1)(\partial^2 f(X_0)/\partial X_n \partial X_2) \dots (\partial^2 f(X_0)/\partial X_n^2)$$

La función $f(X)$ se dice dos veces diferenciable en S , si lo es en cada punto de S .

5) Teorema de Taylor

a) Sean S un conjunto abierto convexo de \mathfrak{R}^n y $f: S \rightarrow \mathfrak{R}$ una función diferenciable en S . Si $X \in S$ y $X_0 \in S$, el desarrollo de Taylor de primer orden, llamado teorema del valor medio, se escribe:

$$f(X) = f(X_0) + [\nabla f(X^*)]^T (X - X_0)$$

donde $X^* = \alpha X_0 + (1 - \alpha) X$ para $\alpha \in (0, 1)$.

b) Dados S un conjunto abierto convexo de \mathfrak{R}^n y $f: S \rightarrow \mathfrak{R}$ dos veces diferenciable en S , si $X \in S$ y $X_0 \in S$, la fórmula de Taylor de segundo orden es :

$$f(X) = f(X_0) + [\nabla f(X_0)]^T (X - X_0) + \frac{1}{2} (X - X_0)^T H(X^*) (X - X_0)$$

donde $H(X^*)$ es la matriz hessiana de f en $X^* = \alpha X_0 + (1 - \alpha) X$ con $\alpha \in (0, 1)$.

B- Conjunto Convexos

1) Definición

Un conjunto S de puntos de \mathfrak{R}^n es denominado conjunto convexo si y solo si para cualesquier par de puntos sea $X_1 \in S$ y $X_2 \in S$, y para cualquier escalar $\alpha \in [0, 1]$, se verifica que $X_3 = \alpha X_2 + (1 - \alpha) X_1 \in S$. El conjunto de puntos X_3 se llama segmento cerrado y la igualdad anterior se denomina combinación lineal convexa de X_1 y X_2 .

2) Hiperplano

a) Siendo b un número real, se define como hiperplano en \mathfrak{R}^n al conjunto de puntos que satisfacen la ecuación lineal $a_1 x_1 + a_2 x_2 + \dots + a_n x_n = b$

Todo hiperplano divide los puntos del espacio \mathfrak{R}^n en dos semiespacios:

$$a_1 x_1 + a_2 x_2 + \dots + a_n x_n \leq b$$

$$a_1 x_1 + a_2 x_2 + \dots + a_n x_n \geq b$$

b) Teorema: Todo semiespacio de \mathfrak{R}^n es un conjunto convexo.

En efecto sean X_1 y X_2 dos puntos pertenecientes al semiespacio $AX \geq b$. Indiquemos con $X = \alpha X_1 + (1 - \alpha) X_2$ un punto arbitrario del segmento definido por X_1 y X_2 . Entonces:

$$AX = A[\alpha X_1 + (1 - \alpha) X_2] = \alpha AX_1 + (1 - \alpha) AX_2 \geq \alpha b + (1 - \alpha) b = b$$

Por lo tanto, también X pertenece al semiespacio considerado, luego $AX \geq b$ es un conjunto convexo.

Análogamente se demuestra que el semiespacio $AX \leq b$ es un conjunto convexo.

c) Corolario: El hiperplano $AX = b$ es un conjunto convexo por ser la intersección de dos conjuntos convexos.

C- Funciones Convexas

1) Definición

Una función $f(X) = f(x_1, x_2, \dots, x_n)$ es una función convexa sobre un conjunto convexo S si y solo si para todo X_1 y $X_2 \in S$ se verifica la desigualdad:



$$f[\alpha X_1 + (1 - \alpha) X_2] \leq \alpha f(X_1) + (1 - \alpha)f(X_2) \quad \forall X_1, X_2 \in S \text{ y } \alpha \in [0, 1]$$

2) Teoremas

a) Sea S un conjunto convexo en \mathfrak{R}^n y $g: S \rightarrow \mathfrak{R}$ una función convexa. Luego el conjunto $T = \{X \in S: g(X) \leq k\}$, donde k es un número real, es un conjunto convexo.

Demostración: En efecto, siendo $X_1, X_2 \in T$, entonces $X_1, X_2 \in S$ y $g(X_1) \leq k$ y $g(X_2) \leq k$.

Ahora, dado $\alpha \in (0, 1)$, resulta $X = \alpha X_1 + (1 - \alpha)X_2 \in S$ por ser S convexo.

Entonces, por la convexidad de $g(X)$ se tiene que $g(X) \leq \alpha g(X_1) + (1 - \alpha)g(X_2) \leq \alpha k + (1 - \alpha)k = k$.

Por lo tanto $X \in T$ y el conjunto T es convexo.

b) Siendo S un conjunto convexo de \mathfrak{R}^n y $g: S \rightarrow \mathfrak{R}$ una función cóncava, el conjunto $T = \{X \in S: g(X) \geq k\}$ es un conjunto convexo, siendo $k \in \mathfrak{R}$.

c) Toda función lineal $h(X)$ definida en un conjunto convexo S de \mathfrak{R}^n es a la vez convexa y cóncava en su dominio.

Demostración: si $h(X)$ es lineal entonces $h(\alpha X) = \alpha h(X)$; $h(X_1 + X_2) = h(X_1) + h(X_2)$. Por lo tanto $h[\alpha X_1 + \beta X_2] = h(\alpha X_1) + h(\beta X_2) = \alpha h(X_1) + \beta h(X_2)$

y siendo $\alpha + \beta = 1$, $\alpha \geq 0$ y $\beta \geq 0$, luego $\beta = 1 - \alpha$, se tiene:

$$h[\alpha X_1 + (1 - \alpha)X_2] = \alpha h(X_1) + (1 - \alpha)h(X_2)$$

3) Funciones Convexas Diferenciables – Teoremas:

a) Sea S un conjunto abierto convexo de \mathfrak{R}^n y $f: S \rightarrow \mathfrak{R}$ una función diferenciable en S . Entonces f es convexa si y solo si para cualquier $X_0 \in S$ se tiene:

$$f(X) \geq f(X_0) + [\nabla f(X_0)]^T (X - X_0) \quad \forall X \in S$$

Si f es convexa en S y si X y $X_0 \in S$, se cumple la desigualdad

$$f[\alpha X + (1 - \alpha)X_0] \leq \alpha f(X) + (1 - \alpha)f(X_0) \quad (1)$$

$\forall \alpha \in [0, 1]$. Por ser f diferenciable en S , se tiene:

$$\begin{aligned} f[\alpha X + (1 - \alpha)X_0] &\leq f[X + \alpha(X - X_0)] = \\ &= f(X_0) + [\nabla f(X_0)]^T \alpha(X - X_0) + \alpha \bullet X - X_0 \bullet \omega[\alpha(X - X_0)] \end{aligned}$$

Reemplazando en (1)

$$\begin{aligned} f(X_0) + [\nabla f(X_0)]^T \alpha(X - X_0) + \alpha \bullet X - X_0 \bullet \omega[\alpha(X - X_0)] &\leq \alpha f(X) + (1 - \alpha)f(X_0) = \\ &= f(X_0) - \alpha[f(X) - f(X_0)] \end{aligned}$$

$$\therefore [\nabla f(X_0)]^T (X - X_0) + \bullet X - X_0 \bullet \omega[\alpha(X - X_0)] \leq f(X) - f(X_0)$$

Entonces cuando $\alpha \rightarrow 0$, queda

$$[\nabla f(X_0)]^T (X - X_0) \leq f(X) - f(X_0) \quad (2)$$

$$\therefore f(X) \geq f(X_0) + [\nabla f(X_0)]^T (X - X_0)$$

Recíprocamente, por cumplirse (2) podemos escribir las desigualdades:

$$f(X_0) - f[\alpha X + (1 - \alpha) X_0] \geq \{\nabla f[\alpha X + (1 - \alpha)X_0]\}^T \alpha(X_0 - X) \quad (3)$$

$$f(X) - f[\alpha X + (1 - \alpha) X_0] \geq \{\nabla f[\alpha X + (1 - \alpha)X_0]\}^T (1 - \alpha)(X - X_0) \quad (4)$$

Multiplicando (3) por $(1 - \alpha)$ y (4) por α , sumando miembro a miembro se tiene:



$$f[\alpha X + (1 - \alpha)X_0] \leq \alpha f(X) + (1 - \alpha)f(X_0)$$

y la función f resulta convexa en S .

b) Sea S un conjunto abierto convexo de \mathfrak{R}^n y $f : S \rightarrow \mathfrak{R}$ dos veces diferenciable en S . Luego f es convexa si y solo si la matriz hessiana es semidefinida positiva en cada punto de S .

Demostración: Suponiendo que f es convexa y sea $X_0 \in S$. Se quiere probar que $X^T H(X_0) X \geq 0$ para cada $X \in \mathfrak{R}^n$. Puesto que S es abierto, para cualquier $X \in \mathfrak{R}^n$, será $X_0 + \alpha X \in S$ para $|\alpha| \neq 0$ suficientemente pequeño. Entonces, por el teorema anterior es:

$$f(X_0 + \alpha X) \geq f(X_0) + \alpha [\nabla f(X_0)]^T X \tag{5}$$

y por ser dos veces diferenciable, se puede escribir:

$$f(X_0 + \alpha X) = f(X_0) + \alpha [\nabla f(X_0)]^T X + \frac{1}{2} \alpha^2 X^T H(X_0) X + \alpha^2 \bullet X \bullet^2 \omega(X_0, \alpha X) \tag{6}$$

Restando (5) - (6), se tiene:

$$\frac{1}{2} \alpha^2 X^T H(X_0) X + \alpha^2 \bullet X \bullet^2 \omega(X_0, \alpha X) \geq 0$$

Dividiendo por α^2 y haciendo $\alpha \rightarrow 0$, resulta:

$$X^T H(X_0) X \geq 0$$

Recíprocamente, suponiendo ahora que la matriz hessiana es semidefinida positiva en cada punto de S . Sean X y $X_0 \in S$, la fórmula de Taylor de segundo orden proporciona:

$$f(X) = f(X_0) + [\nabla f(X_0)]^T (X - X_0) + \frac{1}{2} (X - X_0)^T H(X^*) (X - X_0) \tag{7}$$

siendo $X^* = \alpha X_0 + (1 - \alpha)X$ y $\alpha \in (0, 1)$. En consecuencia, $X^* \in S$ y entonces por hipótesis $H(X^*)$ es semidefinida positiva.

Por lo tanto: $(X - X_0)^T H(X^*) (X - X_0) \geq 0$

Concluyendo que de (7) resulta:

$$f(X) \geq f(X_0) + [\nabla f(X_0)]^T (X - X_0)$$

Ya que la anterior desigualdad es válida para cada X y $X_0 \in S$, f es convexa por el teorema a).

Este teorema es útil para probar la convexidad o concavidad de un función dos veces diferenciable. En especial, si la función es cuadrática, la matriz hessiana es independiente del punto considerado.

c) Sea S un conjunto abierto convexo de \mathfrak{R}^n y $f : S \rightarrow \mathfrak{R}$ dos veces diferenciable en S . Si la matriz hessiana es definida positiva en cada punto de S , entonces f es estrictamente convexa. Recíprocamente, si f es estrictamente convexa, entonces la matriz hessiana es semidefinida positiva en cada punto de S . Sin embargo, si f es estrictamente convexa y cuadrática, la matriz hessiana es definida positiva.



Apartado de Aplicación del Sistema DERIVE for Windows

```
#1: "Programación Cuadrática y Selección de Cartera de Inversión"
#2: "Problema Motivador"
#3: "Matriz U de variancias y covariancias:"
#4:  $v := \begin{bmatrix} 0.0081 & 0.02 \\ 0.02 & 0.0036 \end{bmatrix}$ 
#5: "Menores Principales"
#6:  $v \cdot 11 = 0.0081 > 0$ 
#7:  $DET(v)$ 
#8:  $-0.000370840$ 
#9: "De acuerdo a los menores principales (criterio de Sylvester) la matriz v es indefinida"
#10: "Autovalores"
#11:  $EIGENVALUES(v, \lambda)$ 
#12:  $[\lambda = 0.0259761, \lambda = -0.0142761]$ 
#13: "Los signos de los autovalores corroboran que v es indefinida"
#14: "Por lo tanto, la función objetivo no es convexa como lo requieren las condiciones de"
#15: "programación cuadrática"
```

Bibliografía

- BELLMAN, R. (1960) "Introduction to Matrix Analysis".
- CHIANG, A. C. (1996) "Métodos Fundamentales de Economía Matemática". Edit. Mc Graw Hill. 3° Ed.
- HADLEY, G. (1969) "Álgebra Lineal / Linear Algebra" Fondo Educativo Interamericano. Edición bilingüe.
- MANACORDA, A.E.J. (1995) "Programación No Lineal con Restricciones". Fundación San Cristobal.
- ROSS, S. A.; R. W. WESTERFIELD y J. F. JAFFE (1997) "Finanzas Corporativas". Edit. Mc Graw Hill. 3° Ed.
- WISTON, W. (1994) "Investigación de Operaciones". Ed. Mc Graw Hill.



Indice

Introducción.....	1
I – Problema Motivador.....	2
II – Programación No Lineal con Restricciones	3
III – Programación Convexa y Programación Cuadrática	7
IV – Selección de Carteras de Activos Financieros	11
V – Resolución del Problema Motivador	14
VI – Consideraciones Finales y Conclusiones	18
VII – Apéndice	19
VIII – Apartado de Aplicaciones al Sistema DERIVE for Windows	24
IX – Bibliografía	25
X – Indice	26