

# Notas de Optimización

---

## Introducción a la Optimización Numérica

---

María Cristina Maciel <sup>1</sup>

Córdoba – Junio, 2000

<sup>1</sup>Ph.D in Mathematical Sciences, Rice University, Houston, Texas. Profesora Asociada del Departamento de Matemática de la Universidad Nacional del Sur, Avenida Alem 1253, 8000 Bahía Blanca, ARGENTINA. e-mail: [immaci@criba.edu.ar](mailto:immaci@criba.edu.ar)

## Prólogo

Estas notas corresponden a un breve e introductorio curso sobre Optimización Numérica que dicté en la Facultad de Matemática, Astronomía y Física de la Universidad Nacional de Córdoba.

La audiencia estaba integrada por alumnos y docentes de áreas tan diversas como Matemática, Física y Economía. Dado lo heterogeneo del grupo y limitada por el tiempo, las notas incluyen ideas generales de los temas que cubre el área de Optimización Numérica, y una extensa bibliografía que va desde textos clásicos así como publicaciones muy recientes.

En la primer clase se introduce el problema general de optimización y se presentan algunos ejemplos. En las tres clases siguientes se analiza el problema de optimización irrestricto y el de sistemas de ecuaciones algebraicas no lineales. La referencia básica de estas clases es el tecto de Dennis y Schnabel [12]. Durante la clase 5 se presenta el problema de programación no lineal, estableciendo condiciones necesarias y suficientes para la existencia de solución. Finalmente en la última clase se detalla el bien conocido algoritmos de programación cuadrática sucesiva para resolver el problema de programación no lineal.

Todas las clases incluyen algún ejercicio con la intención de atraer estudiantes que se interesen por un área novedosa, con un número considerable de aplicaciones en el mundo real y tanta investigación que realizar como es la de Optimización Numérica.

Deseo agradecer a Andres Barrea, Rodrigo Burgesser, Pablo Di Ronco, Gabriela Martínez, Fernando Menzaque, Silvina Smith y Germán Torres por su interés en este curso. Quiero expresar un especial agradecimiento a Cristina Turner, quien hizo todo lo posible para que esta visita al FaMAF se concretara y a Elvio Pilotta, por sus observaciones y comentarios sobre estas notas.

# Contenido

<b>1</b>	<b>Problema general de Optimización</b>	<b>4</b>
1.1	Definiciones - Clasificación de problemas . . . . .	4
1.2	Ejemplos de problemas de optimización . . . . .	7
1.2.1	El problema de viajante . . . . .	7
1.2.2	El problema isoperimétrico . . . . .	9
1.2.3	El cociente de Rayleigh . . . . .	10
1.2.4	Ejercicio . . . . .	11
<b>2</b>	<b>Caso irrestricto y problemas relacionados</b>	<b>12</b>
2.1	Ejemplo: problema de estimación de parámetros . . . . .	13
2.2	Problemas relacionados . . . . .	15
2.2.1	Características de los problemas no lineales . . . . .	16
2.3	Sistemas no lineales de ecuaciones algebraicas . . . . .	17
2.3.1	Método de Newton . . . . .	18
2.3.2	Velocidad de convergencia . . . . .	19
2.3.3	Convergencia del método de Newton . . . . .	21
2.3.4	Teorema de Kantorovich . . . . .	23
2.3.5	Teorema de punto fijo . . . . .	24
2.3.6	Ejercicios . . . . .	25
<b>3</b>	<b>Métodos cuasi-Newton para SNL</b>	<b>26</b>
3.1	Método de Newton/diferencias finitas . . . . .	26
3.2	Métodos secantes para sistemas no lineales . . . . .	29
3.2.1	Método de Broyden . . . . .	32
3.2.2	Ejercicios . . . . .	35
<b>4</b>	<b>Minimización sin restricciones</b>	<b>36</b>
4.1	Condiciones de optimalidad . . . . .	36
4.2	Direcciones de descenso . . . . .	38

4.2.1	Método de Newton . . . . .	40
<b>5</b>	<b>Minimización con restricciones</b>	<b>44</b>
5.1	Un ejemplo . . . . .	44
5.2	Multiplicadores de Lagrange . . . . .	45
5.2.1	Geometría del problema . . . . .	46
5.2.2	Existencia de los multiplicadores de Lagrange . . . . .	47
5.3	Condiciones de optimalidad . . . . .	51
5.3.1	Condición necesaria de primer orden . . . . .	51
5.3.2	Condición necesaria de segundo orden . . . . .	51
5.3.3	Condiciones suficientes . . . . .	52
5.3.4	Aplicación . . . . .	53
5.3.5	Regularidad . . . . .	55
<b>6</b>	<b>Algunos métodos numéricos para PNL</b>	<b>56</b>
6.1	Método de Newton . . . . .	57
6.1.1	Ejercicios . . . . .	61
6.2	Programación cuadrática sucesiva . . . . .	61
6.2.1	Construcción del modelo cuadrático . . . . .	62
6.2.2	El algoritmo de programación cuadrática sucesiva . . . . .	64
6.2.3	Solución del problema cuadrático . . . . .	65
6.2.4	Formulación del algoritmo de Dennis y Williamson . . . . .	69
6.2.5	Formulación del algoritmo de Dennis y Maciel . . . . .	72
6.2.6	Ejercicios . . . . .	78

# Clase 1

## Problema general de Optimización

### 1.1 Definiciones - Clasificación de problemas

**Definición 1.1 (Problema general de optimización, PGO)**

*Dado un conjunto  $S$ , y una función  $f : S \rightarrow \mathbb{R}$  el problema de hallar  $x_* \in S$  tal que*

$$f(x_*) \leq f(x), \quad \forall x \in S$$

*se dice problema general de optimización.*

*Se denota*

$$\min_{x \in S} f(x).$$

**Convención:** usualmente se habla de minimizar y no de maximizar. Claramente maximizar  $f$  puede obtenerse de minimizar  $-f$ .

### Clasificación del PGO

Se puede clasificar el problema general de optimización en optimización discreta y optimización continua, analizando el conjunto  $S$ . Más precisamente.

**Definición 1.2 (Optimización discreta)**

*Se dice que PGO es un problema de optimización discreta si  $\text{card}(S) \leq \aleph_0$ , esto es si  $S$  es un conjunto finito o numerable.*

### Ejemplo 1.1

1.

$$S = \{1, 4, -8, 8\}, \quad f(x) = x^2$$

$$x_\star = 1, \quad \min f(x) = f(x_\star) = 1.$$

2.

$$S = \{x \in \mathbb{Z}, x = 2k + 1, k \in \mathbb{Z}\} \quad f(x) = x^2$$

$$x_\star = 1, x_{\star\star} = -1, \quad f(x_\star) = f(x_{\star\star}) = 1.$$

**Observación:**

- Un problema de optimización discreta puede no tener solución.

**Definición 1.3 (Optimización continua)**

Si  $\text{card}(S) > \aleph_0$ , el problema PGO se dice problema continua

- El problema PGO establecido como en 1.2 resulta demasiado general.

**Definición 1.4 (Problema estandar de optimización)**

Sea  $X$  un espacio vectorial real,  $S \subset X$  (no necesariamente un subespacio),  $f : X \rightarrow \mathbb{R}$ . Entonces PGO se dice problema estandar de optimización.

1. Dimensión del problema: es la dimensión del espacio vectorial.
2.  $S = X$ , el problema se dice problema sin restricciones.
3.  $S \subset X$ , el problema se dice problema con restricciones.

**Ejemplo 1.2**

$$f : \mathbb{R}^5 \rightarrow \mathbb{R}, \quad f(x) = x_1 + x_2^2 + \sin x_3 + x_4 x_4.$$

El problema

$$\min_{x \in \mathbb{R}^5} f(x)$$

es un problema sin restricciones de dimensión 5.

Si

$$S = \{x \in \mathbb{R}^5 : x_1 + x_4 x_5 = \sin x_2, x_3 + x_2^2 = 0\}$$

el problema

$$\min_{x \in S} f(x),$$

es un problema con restricciones de dimensión 5.

- Si la dimensión del problema no es muy alta y las funciones involucradas son suficientemente diferenciables con derivadas sencillas de calcular, el problema estandar se puede resolver aplicando las técnicas elementales de cálculo.
- Sin embargo, si el problema es de una dimensión mediana o alta ( $n \geq 50$ ) o bien si las derivadas analíticas no están disponibles se puede aplicar un método numérico para encontrar algún mínimo local. Por ejemplo, el método de Newton.

**Definición 1.5 (Problema general de programación, PGP)**

El problema estandar se dice problema general de programación si el conjunto  $S$  tiene la forma

$$S = \{x \in X : c_i(x) = 0, i \in I, c_j(x) \geq 0, j \in D\},$$

donde  $I, D \subseteq \mathbb{N}$  son conjuntos de índice y  $f, c_i : X \rightarrow \mathbb{R}$ .

Se denota como

$$\begin{aligned} \min \quad & f(x) \\ & c_i(x) = 0, \quad i \in I \\ & c_j(x) \geq 0, \quad j \in D. \end{aligned}$$

**Ejemplo 1.3**

$$\begin{aligned} \min \quad & x_1 + x_2^2 + \sin x_3 + x_4 x_4 \\ & x_1 + x_4 x_5 - \sin x_2 = 0, \\ & x_3 + x_4^2 = 0. \end{aligned}$$

$$I = \{1, 2\}, \quad D = \emptyset.$$

**Clasificación del PGP**

Vamos a clasificar el problema general de programación observando el espacio vectorial  $X$  y los conjuntos de índices  $I$  y  $D$ .

1. Si  $\dim X = \infty$  y  $\text{card}(I) = \text{card}(D) = \infty$ , el problema se dice infinito.
2. Si  $\dim X = \infty$  y  $\text{card}(I)$  y  $\text{card}(D)$  son ambos finitos, el problema se dice semi-finito.
3. Si  $\dim X < \infty$  y  $\text{card}(I)$  o bien  $\text{card}(D)$  es infinito, el problema se dice semi-infinito.

4. Si  $\dim X < \infty$  y  $\text{card}(I)$  y  $\text{card}(D)$  son ambos finitos, el problema se dice finito. En particular el problema finito se puede escribir como

$$\begin{aligned} \min \quad & f(x) \\ \text{s.a} \quad & \\ & c_i(x) = 0 \quad i \in I = \{1, \dots, m\} \\ & c_i(x) \geq 0 \quad i \in D = \{m+1, \dots, m+p\}. \end{aligned}$$

con  $f, c_i : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $i \in I \cup D$ .

Si para algún  $i$ ,  $c_i(x)$  o  $f(x)$  es no lineal el problema de optimización se dice de **programación no lineal, (PNL)**.

### Casos especiales:

- **Programación cuadrática (PC)**

$$f(x) = \frac{1}{2}x^T Hx + g^T x + \alpha$$

con  $H \in \mathbb{R}^{n \times n}$ ,  $H = H^T$ ,  $g \in \mathbb{R}^n$ ,  $\alpha \in \mathbb{R}$  y

$$c_i(x) = a_i^T x + \alpha_i, \tag{1.1}$$

con  $a_i \in \mathbb{R}^n$ ,  $\alpha_i \in \mathbb{R}$  para todo  $i \in I \cup D$ .

- **Programación lineal (PL)**

$$f(x) = c^T x + \alpha$$

con  $c \in \mathbb{R}^n$ ,  $\alpha \in \mathbb{R}$  y  $c_i$  definidas como en (1.1).

Los interesados en temas de programación lineal son remitidos a [6, 25, 28, 35].

- **Programación no diferenciable y programación convexa** se puede consultar en [4, 16, 30].

## 1.2 Ejemplos de problemas de optimización

### 1.2.1 El problema de viajante

Un viajante debe visitar  $n$ -ciudades distintas, comenzando y terminando su viaje en la misma ciudad. Cada ciudad puede ser visitada solamente una vez.

**PROBLEMA:** *Encontrar un viaje de longitud mínima.*

Sea  $\mathcal{C} = \{1, 2, \dots, n\}$ , el conjunto de ciudades.

Supongamos que el viaje comienza y termina en la ciudad 1.

Definimos



$d_{ij}$ , la distancia entre las ciudades  $i$  y  $j$ . ( $d_{ii} = 0$ ),

$x_{ij}$ , representa el viaje desde  $i$  a  $j$ ,

$$x_{ij} = \begin{cases} 1 & \text{si el recorrido incluye viaje de } i \text{ a } j \\ 0 & \text{otro caso.} \end{cases}$$

Cada recorrido implica *partir* de la ciudad  $i$ , para cada  $i$ .

$$\sum_{j=1}^n x_{ij} = 1, \quad i = 1, \dots, n. \quad (1.2)$$

Cada recorrido implica *llegar* a la ciudad  $j$ , para cada  $j$ .

$$\sum_{i=1}^n x_{ij} = 1, \quad j = 1, \dots, n. \quad (1.3)$$

La longitud del viaje es:

$$f(x_{11}, x_{12}, \dots, x_{ij}, \dots, x_{n-1,n}) = \sum_{i=1}^n \sum_{j=1}^n d_{ij} x_{ij}. \quad (1.4)$$

El problema es

$$\begin{aligned} \min \quad & \sum_{i=1}^n \sum_{j=1}^n d_{ij} x_{ij} \\ \text{s.a} \quad & \\ & \sum_{j=1}^n x_{ij} = 1, \quad i = 1, \dots, n \\ & \sum_{i=1}^n x_{ij} = 1, \quad j = 1, \dots, n \\ & x_{ij} \in \{0, 1\}. \end{aligned}$$

El problema del viajante es un problema discreto.  $\text{card}(S) = (n - 1)!$ .

Este es un ejemplo de una clase de problemas que se conoce con el nombre de programación entera. Es importante observar que se están usando variables 0 – 1 para representar las posibles elecciones de los recorridos. En general, se considera que un evento puede o no ocurrir y es parte del problema decidir entre esas dos posibilidades. Para modelar esta dicotomía se usa una variable binaria definida así

$$x = \begin{cases} 1 & \text{si el evento ocurre} \\ 0 & \text{si el evento no ocurre.} \end{cases}$$

Otros problemas famosos que necesitan de variables binarias son:

- El problema de la mochila.
- Problemas de asignación.
- Problemas de apareamiento.

Para los interesados en optimización discreta un clásico es el texto de Nemhauser y Wolsey [26].

### 1.2.2 El problema isoperimétrico

**PROBLEMA:** Hallar una curva  $f : [t_0, t_1] \rightarrow \mathbb{R}^n$ , tal que para todo  $t \in [t_0, t_1]$ ,  $f(t) = (x_1(t), \dots, x_n(t))^T$  minimiza la funcional

$$J(x_1, \dots, x_n) = \int_{t_0}^{t_1} F(t, x_1, \dots, x_n, \dot{x}_1, \dots, \dot{x}_n) dt,$$

bajo las condiciones:

$$f(t_0) = f_0, \quad f(t_1) = f_1, \quad (\text{condiciones de borde})$$

$$\int_{t_0}^{t_1} F_i(t, x_1, \dots, x_n, \dot{x}_1, \dots, \dot{x}_n) dt = l_i, \quad 1 \leq i \leq m$$

donde  $l_i$  es constante y  $m \leq n$

El problema isoperimétrico más simple puede ser establecido como sigue.

*Entre todas las posibles curvas planas, cerradas, simples y de perímetro dado, encontrar aquella que encierra mayor área.*

Sea  $\gamma$  una curva de longitud  $\ell$ . Supongamos que  $\gamma = \mathfrak{Sm}(f)$  donde  $f : [t_0, t_1] \rightarrow \mathbb{R}^2$  tal que  $f(t) = (x_1(t), x_2(t))^T$  satisface

- (1)  $f$  tiene derivada primera continua
- (2)  $\dot{x}_1^2 + \dot{x}_2^2 \neq 0$
- (3)  $x_1(t_0) = x_1(t_1), \quad x_2(t_0) = x_2(t_1)$
- (4) la longitud de  $\gamma$  esta dada por

$$\ell = \int_{t_0}^{t_1} \sqrt{\dot{x}_1^2 + \dot{x}_2^2} dt.$$

Sea

$$S = \{f : [t_0, t_1] \rightarrow \mathbb{R}^2 : f \text{ satisface las condiciones (1) – (4)}\}.$$

$$S \subset \mathcal{C}^1([t_0, t_1]) = \{f : [t_0, t_1] \rightarrow \mathbb{R}^2 : f' \text{ continua}\}.$$

Hay que hallar la curva de longitud  $\ell$  que encierra la mayor superficie. Esto significa que se debe maximizar

$$\begin{aligned} J(f) &= \frac{1}{2} \oint_{\gamma} x_1 dx_2 - x_2 dx_1 \\ &= \frac{1}{2} \int_{t_0}^{t_1} [x_1 \dot{x}_2 - x_2 \dot{x}_1] dt. \end{aligned}$$

Por lo tanto el problema queda establecido como

$$\max_{f \in S} J(f).$$

La solución es la circunferencia de perímetro  $\ell$ , cuya parametrización es

$$f_{\star}(t) = \left( \frac{\ell}{2\pi} \cos 2\pi \frac{t - t_0}{t_1 - t_0}, \frac{\ell}{2\pi} \sin 2\pi \frac{t - t_0}{t_1 - t_0} \right)^T.$$

En este ejemplo,  $\text{card}(S) > \aleph_0$ , entonces el problema es continuo y como  $\dim_{\mathbb{R}} \mathcal{C}^1([t_0, t_1]) = \infty$ , el problema es infinito.

Este tipo de problema corresponden al área de cálculo de variaciones. Otros ejemplos clásicos que pueden plantearse en forma similar se pueden mencionar:

El problema del *spline* cúbico.

El problema de hallar el camino más corto entre dos puntos.

El problema de la braquistocrona.

La herramienta para resolver estos problemas son las ecuaciones de Euler-Lagrange y la desigualdad variacional.

### 1.2.3 El cociente de Rayleigh

**PROBLEMA:** Encontrar el mayor autovalor de una matriz  $A \in \mathbb{R}^{n \times n}$ , simétrica y definida positiva.

Como  $A$  es simétrica y definida positiva todos sus autovalores son números reales estrictamente positivos. Entonces  $\lambda \in \mathbb{R}$  es autovalor de  $A$  y  $v \in \mathbb{R}^n$ ,  $v \neq 0$  su autovector asociado si

$$Av = \lambda v$$

de donde

$$v^T A v = \lambda v^T v.$$

Por lo tanto

$$\lambda = f(v) = \frac{v^T A v}{v^T v} > 0 \quad (1.5)$$

que es el **cociente de Rayleigh**. Conocido los autovectores de  $A$ , (1.5) permite hallar los correspondientes autovalores. Entonces si definimos

$$S = \{v \in \mathbb{R}^n : v \neq 0\} \subset \mathbb{R}^n,$$

el problema de hallar el mayor autovalor de una matriz simétrica, definida positiva se puede establecer como sigue

$$\max_{v \in S} \frac{v^T A v}{v^T v}.$$

Como  $\dim \mathbb{R}^n = n$  y  $\text{card}(S) > \aleph_0$  el problema es continuo y finito.

## 1.2.4 Ejercicio

### Ejercicio 1.1 Regresión lineal

Sea el siguiente conjunto de puntos del plano

$$\{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$$

y considere el problema de ajustar estos datos mediante una función lineal,  $f(x) = ax + b$ . Establezca este problema usando las normas  $\ell_p$  con  $p = 1, 2, \infty$ . Verifique que si  $p = 1$  o  $\infty$  el problema se puede plantear como un problema de programación lineal, mientras que si  $p = 2$  el problema es de programación cuadrática.

Una cantidad considerable de distintos tipos de problemas de optimización se pueden hallar en los textos de [2, 3] y [23].

## Clase 2

# Caso irrestricto y problemas relacionados

El problema que vamos a comenzar a estudiar en esta clase es el de hallar los puntos que minimizan a una función definida en  $\mathbb{R}^n$ . Esto es dada  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , hallar  $x_* \in \mathbb{R}^n$  tal que

$$f(x_*) \leq f(x), \quad \forall x \in \mathbb{R}^n$$

**Notación:**

$$\min_{x \in \mathbb{R}^n} f(x)$$

*¿Por qué interesa este problema ? .*

1. Problemas CON restricciones aparecen en muchas aplicaciones. Algunos de estos problemas pueden transformarse en problemas sin restricciones.
2. También ocurre que algunos métodos para resolver problema con restricciones consisten en resolver una sucesión de problemas sin restricciones (por ejemplo, los llamados métodos de penalización).  
Por lo tanto, es claro que estudiar el problema sin restricciones y los métodos que lo resuelven es un problema central en el desarrollo de algoritmos para problemas con restricciones.
3. Por otra parte en muchas aplicaciones, se presentan problemas sin restricciones en su forma pura.

## 2.1 Ejemplo: problema de estimación de parámetros

Supongamos que tenemos una variable  $y$ , (variable dependiente) y variables  $x_1, \dots, x_p$ , (variables independientes).

Supongamos que se han observado valores  $y_1, \dots, y_r$  de  $y$ , asociados a valores  $x_{ij}, i = 1, \dots, r, j = 1, \dots, p$  de las variables  $x_j, j = 1, \dots, p$ .

**PROBLEMA:** establecer la relación entre las variables  $y$  y las variables  $x_1, \dots, x_p$ .

- Este problema aparece en distintas áreas como economía, medicina, biología, ingeniería, etc.
- Se desea saber la relación exacta entre las variables para poder predecir un comportamiento que no se puede observar. También, a partir de esa relación se desea controlar los futuros valores de  $y$  modificando los valores de  $x_j, j = 1, \dots, p$ .

$$y = \mathbf{F}(x_1, \dots, x_p; b_1, \dots)$$

$\mathbf{F}$  es la funcional y  $b_1, b_2, \dots$  son los parámetros a estimar.

- Los pasos para resolver este problema de estimación de parámetros son:
  1. Formular un modelo cualitativo de causa y efecto identificando las variables independientes.
  2. Obtener una colección de datos.
  3. *Formular el modelo matemático postulando la forma de la funcional que relaciona los datos*
  4. *Determinar los valores de los parámetros de la funcional que mejor expliquen los datos observados.*
  5. Obtener intervalos de confianza para los mejores valores de los parámetros.

En los modelos de regresión que se originan en ciencias sociales, el paso 3. involucra la formulación de una funcional lineal o logarítmica y el paso 4. involucra la resolución de un problema de cuadrados mínimos lineales (CML). Precisamente el problema CML es un ejemplo de un problema de minimización sin restricciones.

Supongamos que la relación que se busca es de la forma

$$y = b_{p+1} + \sum_{j=1}^p b_j x_j. \quad (2.1)$$

En este caso, el valor de los parámetros  $b_j$  que mejor explican los datos observados son aquellos para los cuales la suma de los cuadrados del error entre los valores predichos

y los valores reales de la variable dependiente es mínimo. Esto es, aquellos valores  $b_1^*, b_2^*, \dots, b_p^*, b_{p+1}^*$ , tal que minimicen

$$L(b) = L(b_1, \dots, b_p, b_{p+1}) = \sum_{i=1}^r \left( y_i - (b_{p+1} + \sum_{j=1}^p b_j x_{ij}) \right)^2.$$

Las siguientes son algunas razones de la popularidad de la funcional  $L$ .

1. Existe toda una teoría estadística concerniente a  $b^*$ . Esa teoría garantiza que  $b^*$  es un buen estimador de los verdaderos parámetros. Tratando a  $y$  como una variable aleatoria, bajo ciertas hipótesis sobre  $L$  se puede probar que  $b_j^*$ , ( $j = 1, \dots, p+1$ ) son únicos y que tienen una distribución normal. A partir de esa teoría se pueden obtener los intervalos de confianza para los verdaderos valores de  $b_j$ .
2. La segunda razón de la popularidad de  $L$  es que su solución se puede hallar en forma analítica y también existen gran cantidad de códigos para computar la solución y los intervalos de confianza.

Si se define

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_r \end{pmatrix} \in \mathbb{R}^r, \quad A = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1p} \\ \vdots & \vdots & & \vdots \\ 1 & x_{r1} & \cdots & x_{rp} \end{pmatrix} \in \mathbb{R}^{r \times (p+1)},$$

el problema se escribe como

$$\min_{b \in \mathbb{R}^{p+1}} \|Y - Ab\|_2^2$$

y su solución que se puede deducir de las ecuaciones normales es  $b^* = (A^T A)^{-1} A^T Y = A^+ Y$ , siendo  $A^+$  la matriz pseudoinversa de  $A$ .

Otra funcional frecuentemente usada es

$$y = f(a_1, \dots, a_{p+1}) = \exp \left( a_{p+1} \prod_{j=1}^p x_j^{a_j} \right) \quad (2.2)$$

$$\ln y = a_{p+1} + \sum_{j=1}^p a_j \ln x_j.$$

Es claro que los valores que mejor estiman a los parámetros  $a_j$ ,  $j = 1, \dots, p+1$  son encontrados minimizando

$$\sum_{i=1}^r [\ln y_i - (a_{p+1} + \sum_{j=1}^p a_j \ln x_{ij})]^2.$$

Esta tiene una forma similar a la forma anterior. Se pueden escribir las ecuaciones normales cambiando  $\ln y$  por  $y_i$  y  $\ln x_{ij}$  por  $x_{ij}$ .

En ciencias físicas muy rara vez se encuentra un problema de regresión con una funcional lineal o logarítmica. Usualmente la funcional se construye a partir de consideraciones acerca del proceso físico de causa y efecto.

El problema general de regresión tiene la forma

$$\min_{b \in \mathbb{R}^p} \sum_{i=1}^r [y_i - \mathbf{F}(b, x_i)]^2 \quad (2.3)$$

donde  $F(b, x)$  es una función que involucra el vector de parámetros  $b$  y las variables independientes  $x_i$ ,  $i = 1, \dots, r$ .

Es claro que resolver (2.3) no es tan simple como en el caso lineal.

Más ejemplos de problemas del mundo real que se pueden modelar como un problema de optimización no lineal sin restricciones se pueden encontrar en el trabajo de Moré [24]. Este trabajo contiene un conjunto de 14 problemas planteados y casi en todos ellos, el método sugerido para su resolución es algunos de los que se mencionarán en este curso.

## 2.2 Problemas relacionados

El problema de

### I. Minimización sin restricciones (MSR)

$$\min_{x \in \mathbb{R}^n} f(x)$$

está relacionado con otros problemas tales como,

### II. Sistemas de ecuaciones algebraicas no lineales (SNL)

Dado  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ , hallar  $x_* \in \mathbb{R}^n$  tal que  $F(x_*) = 0$ .

### III. Cuadros mínimos no lineales (CMNL)

Dada  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ,  $m \geq n$ , hallar  $x_* \in \mathbb{R}^n$  tal que resuelva,

$$\min_{x \in \mathbb{R}^n} \|F(x)\|_2^2 = \min_{x \in \mathbb{R}^n} \sum_{i=1}^m f_i(x)^2,$$

donde  $F(x) = (f_1(x), \dots, f_m(x))^T$ .

los problemas **II.** y **III.**, bajo hipótesis razonables se pueden tratar como un problema MSR.



- Resolver el sistema no lineal  $F(x) = 0$ , es equivalente a resolver el siguiente problema

$$\min \sum_{i=1}^n f_i(x)^2.$$

**Ejemplo 2.1** El sistema

$$F(x_1, x_2) = \begin{pmatrix} x_1^2 + x_2^3 + 7 \\ x_1 + x_2 + 1 \end{pmatrix},$$

tiene una solución  $x_* = (1, -2)^T$ .

Es claro que  $x_*$  es una solución del siguiente problema de optimización sin restricciones:

$$\min f(x_1, x_2) = (x_1 + x_2^3 + 7)^2 + (x_1 + x_2 + 1)^2.$$

Esto sugiere que las técnicas para resolver uno u otro problema están relacionadas.

- El problema de CMNL es un caso especial de MSR. Sin embargo, no se utilizan los mismos métodos. La idea es modificar las técnicas para resolver el problema MSR de modo de tomar ventajas de la estructura de los problemas y obtener un método mejor.

Estos problemas, son resueltos utilizando **métodos iterativos**. Es decir, por medio de algoritmos que a partir de un punto  $x_0$ , el punto inicial, generan una sucesión, la sucesión de iterados,  $\{x_k\} \subset \mathbb{R}^n$ , que converge a una solución del problema.

Los algoritmos son codificados y la respuesta obtenida rara vez coincide con  $x_*$ . La solución computada,  $\tilde{x}$ , en el mejor de los casos es una aproximación de  $x_*$ .

### 2.2.1 Características de los problemas no lineales

De acuerdo a Dennis-Schnabel [12] las siguientes son características, que deben tenerse en cuenta para desarrollar un algoritmo que resuelva un problema no lineal.

1. **Tamaño:** es un concepto que depende del procesador. Un problema se considera pequeño si tiene hasta 100 variables, mientras que si tiene entre 100 y 1000 variables se puede considerar mediano. Finalmente problemas grandes serán aquellos de más de 1000 variables. Claramente esta noción cambia a medida que cambia la tecnología.  
Para problemas de gran tamaño existen algoritmos especiales que explotan la estructura del problema.
2. **Disponibilidad de la derivadas:** muchas veces se sabe que las funciones que intervienen en el problema son continuamente diferenciables. Sin embargo las

derivadas analíticas no están disponibles o son costosas de calcular. Para eso necesario desarrollar algoritmos que trabajen en forma eficiente ante la ausencia de derivadas.

3. **Eficiencia:** ¿qué significa que un algoritmo sea eficiente?. ¡Que sea económico, rápido y convergente !!.  
Problemas costosos de resolver son aquellos donde las funciones que intervienen necesitan mucho tiempo de máquina para ser evaluadas y tal vez lugar de memoria para almacenar cálculos intermedios.  
También puede suceder que para la resolución del problema se necesite resolver subproblemas sencillos relacionados con él.  
Por lo tanto se necesitan desarrollar algoritmos que requieran pocas evaluaciones de funciones y derivadas y que muestren rápida velocidad de convergencia.
4. **Precisión:** dependiendo de la naturaleza del problema se espera que la solución computada tenga unos pocos dígitos de precisión. En general se requieren más dígitos que los se necesitan para asegurar convergencia del algoritmo, pero el punto es que la precisión requerida rara vez está cerca de la precisión de la máquina.
5. **Escalamiento:** Con el calificativo de *problemas pobremente escalados* se quiere significar que los tamaños de las variables difieren considerablemente entre si. Si se ignora este fenómeno el comportamiento de un algoritmo para problemas no lineales se puede ver realmente afectado.

Entonces, tamaño del problema, eficiencia, precisión en la solución y escalamiento del problema son características que deben ser tenidas en cuenta en el desarrollo de un algoritmo que resuelve un problema no lineal, en particular los tres problemas mencionados al principio.

## 2.3 Sistemas no lineales de ecuaciones algebraicas

Consideremos el problema de hallar  $x_* \in \mathbb{R}^n$  tal que  $F(x_*) = 0$ , donde  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ , esto es

$$F(x_1, \dots, x_n) = \begin{pmatrix} f_1(x_1, \dots, x_n) \\ f_2(x_1, \dots, x_n) \\ \vdots \\ f_n(x_1, \dots, x_n) \end{pmatrix}.$$

Comenzamos asumiendo que  $F$  es continuamente diferenciable en  $x \in \mathbb{R}^n$ . Recordemos de los cursos de cálculo en varias variables que en ese caso cada componente  $f_i$  es continuamente diferenciable en  $x$  y la derivada de  $F$  en  $x$  se define como sigue

**Definición 2.1 (Matriz Jacobiana)**

$$F'(x) = \nabla F(x)^T = J(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \dots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \dots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \dots & \frac{\partial f_n}{\partial x_n} \end{pmatrix} \in \mathbb{R}^{n \times n}.$$

**2.3.1 Método de Newton**

Como en el caso unidimensional, el método por elección para resolver el problema presentado arriba es el método de Newton.

Dada una aproximación actual de  $x_*$ , digamos  $x_c$ , se considera un modelo afín de  $F(x)$  en un entorno de  $x_c$ :

$$F(x) = F(x_c + s) = \underbrace{F(x_c) + F'(x_c)s}_{M_c(s)} + R,$$

es decir

$$M_c(s) = F(x_c) + F'(x_c)s \tag{2.4}$$

$$s = x - x_c. \tag{2.5}$$

Hallamos el paso  $s$ , **el paso de Newton**, de modo que  $M_c(s) = 0$ , entonces

$$\begin{aligned} F'(x_c)s &= -F(x_c) \\ s &= -(F'(x_c))^{-1}F(x_c). \end{aligned}$$

Esta construcción nos permite establecer el siguiente algoritmo :

**Algoritmo 2.1 (Método de Newton)**

Sea  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $x_0 \in \mathbb{R}^n$  un punto inicial,

**For**  $k = 0, \dots, \dots$ , *hasta convergencia* **do**

- Resolver el sistema lineal  $F'(x_k)s = -F(x_k)$ .
- Definir  $x_{k+1} = x_k + s_k$ .

**end**

Veamos que significa la expresión *hasta convergencia* en este algoritmo. Un buen criterio de parada del algoritmo es preguntar en cada iteración si  $\|F(x_k)\| \leq \varepsilon_1$  o bien si dos iterados consecutivos satisfacen  $\|x_{k+1} - x_k\| \leq \varepsilon_2 \|x_{k+1}\|$ . También, como en el caso unidimensional hay que fijar el número máximo de iteraciones, para detener el proceso en caso que no haya suficiente progreso hacia  $x_*$ .

En práctica se utiliza  $\|\cdot\|_2$  o  $\|\cdot\|_\infty$ .

**Ejemplo 2.2** Ver [12].

Consideremos

$$F(x_1, x_2) = \begin{pmatrix} x_1^2 + x_2^2 - 2 \\ e^{x_1-1} + x_2^3 - 2 \end{pmatrix}.$$

Una solución es  $x_* = (1, 1)^T$ . Aplicamos el algoritmo 4.2 comenzando desde  $x_0 = (2, 3)$ ,

$$\begin{array}{ll} x_0 = \begin{pmatrix} 2 \\ 3 \end{pmatrix} & e_0 = \|x_0 - x_*\|_\infty = 2 \\ x_1 = \begin{pmatrix} 0.575 \\ 2.117 \end{pmatrix} & e_1 = \|x_1 - x_*\|_\infty = 1.117 \\ x_2 = \begin{pmatrix} 0.312 \\ 1.524 \end{pmatrix} & e_2 = \|x_2 - x_*\|_\infty = 0.688 \approx 7 \times 10^{-1} \\ x_3 = \begin{pmatrix} 1.484 \\ 1.146 \end{pmatrix} & e_3 = \|x_3 - x_*\|_\infty = 0.484 \approx 5 \times 10^{-1} \\ x_4 = \begin{pmatrix} 1.059 \\ 1.034 \end{pmatrix} & e_4 = \|x_4 - x_*\|_\infty = 0.059 \approx 6 \times 10^{-2} \\ x_5 = \begin{pmatrix} 1.0008 \\ 1.0010 \end{pmatrix} & e_5 = \|x_5 - x_*\|_\infty = 10^{-3} \\ x_6 = \begin{pmatrix} 0.9999987 \\ 1.0000020 \end{pmatrix} & e_6 = \|x_6 - x_*\|_\infty = 0.000002 \approx 2 \times 10^{-6} \\ x_7 = \begin{pmatrix} 0.9999999950 \\ 1.00000000009 \end{pmatrix} & e_7 = \|x_7 - x_*\|_\infty = 9 \times 10^{-12}. \end{array}$$

Se observa que la convergencia es  $q$ -cuadrática.

### 2.3.2 Velocidad de convergencia

Supongamos que  $\{x_k\} \subset \mathbb{R}^n$  es una sucesión que converge a un punto  $x_* \in \mathbb{R}^n$ . Esto es

$$\lim_{k \rightarrow \infty} \|x_k - x_*\| = 0.$$

donde  $\|\cdot\|$  es una norma en  $\mathbb{R}^n$ .

**Definición 2.2 (Convergencia  $q$ -lineal)**

Se dice que  $\{x_k\}$  converge  $q$ -linealmente a  $x_*$  si existe una constante  $c \in [0, 1)$  y un índice  $K > 0$  tal que para todo  $k \geq K$  se cumple

$$\|x_{k+1} - x_*\| \leq c\|x_k - x_*\|.$$

Este tipo de convergencia puede no ser satisfactoria en práctica. ¿por qué?

**Definición 2.3 (Convergencia  $q$ -orden  $p$ )**

Se dice que  $\{x_k\}$  converge a  $x_*$  con  $q$ -orden de convergencia  $p$  si existe una constante  $c \geq 0$  y un índice  $K > 0$  tal que para todo  $k \geq K$  se cumple

$$\|x_{k+1} - x_*\| \leq c\|x_k - x_*\|^p.$$

Si  $p = 2$  se tiene convergencia  $q$ -cuadrática y si  $p = 3$  se tiene convergencia  $q$ -cúbica.

**Definición 2.4 (Convergencia  $q$ -superlineal)**

Se dice que  $\{x_k\}$  converge a  $x_*$  con  $q$ -superlinealmente si existe una sucesión  $\{c_k\} \subset \mathbb{R}$  tal que  $c_k \rightarrow 0$  tal que

$$\|x_{k+1} - x_*\| \leq c_k\|x_k - x_*\|.$$

Dicho de otro modo

$$\lim_{k \rightarrow \infty} \frac{\|e_{k+1}\|}{\|e_k\|} = 0.$$

**Ejemplo 2.3** Analicemos que velocidad de convergencia muestran las siguientes sucesiones:

1)  $x_k = 1 + 2^{-k}$

$$\lim_{k \rightarrow \infty} 1 + 2^{-k} = 1 = x_*.$$

Calculamos:

$$|x_{k+1} - 1| = |2^{-(k+1)}| = |2^{-1}2^{-k}| = \frac{1}{2}|2^{-k} + 1 - 1| = \frac{1}{2}|x_k - 1|,$$

es decir que  $\{x_k\}$  converge a 1  $q$ -linealmente con  $c = \frac{1}{2}$ .

2)  $x_k = 1 + 2^{-2^k}$

$$\lim_{k \rightarrow \infty} 1 + 2^{-2^k} = 1 = x_*.$$

Calculamos:

$$|x_{k+1} - 1| = |2^{-2^{k+1}}| = |2^{-(2^k 2)}| = |(2^{-2^k})^2| = |2^{-2^k} + 1 - 1|^2 = \frac{1}{2}|x_k - 1|^2.$$

Luego la convergencia es  $q$ -cuadrática con  $c = 1$ .

El prefijo  $q$  (por *quotient*) en las definiciones de arriba se utiliza para diferenciar este tipo de convergencia de las de tipo  $r$  (por *root*).

**Definición 2.5 (Convergencia  $r$ -orden  $p$ )**

Se dice que  $\{x_k\}$  converge a  $x_*$  con  $r$ -orden de convergencia  $p$  si existe una sucesión  $\{a_k\} \subset \mathbb{R}$  que converge exhibiendo un comportamiento  $q$  de orden  $p$  tal que el error

$$\|x_k - x_*\| = \|e_k\| \leq |a_k|.$$

Para aquellos lectores interesados, la referencia obligada para estudiar los distintos tipos de convergencia es el capítulo 9 de [27].

### 2.3.3 Convergencia del método de Newton

En el método de Newton,  $F(x_c + s)$  es modelada por un modelo afín. Antes de enunciar el teorema de convergencia queremos hallar una cota para la diferencia entre  $F(x_c + s)$  y el modelo  $M_c(s)$ . Para eso necesitamos una hipótesis de continuidad sobre la matriz  $F'(x)$ .

**Definición 2.6 (Continuidad de Lipschitz )**

Sea  $G : \mathbb{R}^n \rightarrow \mathbb{R}^{m \times n}$ ,  $x \in \mathbb{R}^n$ . Sean  $\|\cdot\|$  y  $\|\cdot\|$  normas en  $\mathbb{R}^n$  y  $\mathbb{R}^{m \times n}$  respectivamente. El operador  $G$  se dice Lipschitz continuo en  $x \in \mathbb{D}$  con  $\mathbb{D} \subset \mathbb{R}^n$  abierto, si existe una constante  $\gamma > 0$  tal que para todo  $y \in \mathbb{D}$

$$\|G(y) - G(x)\| \leq \gamma \|x - y\|.$$

La constante  $\gamma$  se dice la constante de Lipschitz.

**Notación:**  $G \in \mathcal{L}ip_\gamma(\mathbb{D})$ .

**Lema 2.1** Sea  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ , continuamente diferenciable en un conjunto abierto y convexo  $\mathbb{D} \subset \mathbb{R}^n$ . Sea  $J \in \mathcal{L}ip_\gamma(\mathbb{D})$  y considerando una norma vectorial y la correspondiente norma inducida, para todo  $x + s \in \mathbb{D}$  se tiene

$$\|F(x + s) - F(x) - J(x)s\| \leq \frac{\gamma}{2} \|s\|^2.$$

**Demostración:**

$$\begin{aligned} F(x + s) - F(x) - J(x)s &= \int_0^1 J(x + st)s dt - J(x)s \\ &= \int_0^1 [J(x + st)s - J(x)s] dt. \end{aligned}$$

Tomando norma, usando propiedades de la integral, de la norma y del hecho que  $J(x)$  es Lipschitz continua con constante  $\gamma$  en  $\mathbb{D}$  se tiene

$$\begin{aligned} \|F(x+s) - F(x) - J(x)s\| &= \left\| \int_0^1 [J(x+st)s - J(x)s] dt \right\| \\ &\leq \int_0^1 \|J(x+st) - J(x)\| \|s\| dt \\ &\leq \gamma \int_0^1 \|ts\| \|s\| dt \\ &\leq \frac{\gamma}{2} \|s\|^2. \end{aligned}$$

■

**Teorema 2.1 (Convergencia del método de Newton)**

Sea  $F : \mathbb{D} \rightarrow \mathbb{R}^n$ ,  $\mathbb{D} \subset \mathbb{R}^n$  abierto y convexo. Asumiendo,

(N1) Existe  $x_* \in \mathbb{D}$  tal que  $F(x_*) = 0$ .

(N2) Existe  $\beta > 0$  tal que  $\|J(x)^{-1}\| \leq \beta$  para todo  $x \in \mathbb{D}$ .

(N3)  $F' \in \mathcal{L}ip_\gamma(\mathbb{D})$ .

Entonces, si existe  $\eta > 0$  tal que para todo  $x_0 \in \mathbb{D}$  tal que  $\|x_0 - x_*\| < \eta$ , la sucesión  $\{x_k\}$  generada por algoritmo 4.2 está bien definida y converge a  $x_*$ .

Además,

$$\|x_{k+1} - x_*\| \leq \beta\gamma \|x_k - x_*\|^2.$$

**Demostración:** Ver [12], Capítulo 5. ■

**Significado de la constante  $\beta\gamma$ :**

De la demostración del teorema se puede observar que el radio de la región de convergencia del método satisface

$$\varepsilon \leq \frac{1}{2\beta\gamma}. \quad (2.6)$$

La constante  $\beta\gamma$  es una medida de la no linealidad relativa de  $F$  en  $x_*$ . Esto se deduce de

$$\frac{\|J(x) - J(x_*)\|}{\|J(x_*)\|} \leq \|J(x_*)^{-1}\| \|J(x) - J(x_*)\| \leq \beta\gamma \|x - x_*\|.$$

- El teorema de convergencia nos dice que el radio de la región de convergencia es inversamente proporcional a la no linealidad relativa de  $F$  en  $x_*$ .

- En práctica la igualdad en (2.6) es una estimación del peor caso que muestra cuán lejos de  $x_*$  se extiende la región de convergencia del método. Si  $F$  es *muy no lineal* el radio será pequeño, mientras que si  $F$  es casi lineal la región de convergencia puede ser más grande.

### Ventajas del método de Newton

- 1) Rápida convergencia local si  $x_0$  es elegido adecuadamente y  $J(x_*)$  es no singular. Si  $J(x_*)$  es singular la convergencia es  $q$ -lineal.
- 2) Si  $F(x) = Ax - b$  el método converge a  $x_*$  en una iteración.

### Desventajas del método de Newton

- 1) El método no converge globalmente.
- 2) En cada iteración necesita calcular la matriz  $J(x)$ .
- 3) En cada iteración necesita resolver un sistema lineal. Luego el trabajo aritmético es  $\mathcal{O}(n^3)$  por iteración, más el inconveniente de posible mal condicionamiento.

#### 2.3.4 Teorema de Kantorovich

##### Teorema 2.2 (Teorema de Kantorovich)

Sea  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $x_0 \in \mathbb{R}^n$ ,  $r > 0$ ,  $F \in C^1$  en  $\mathcal{N}(x_0, r)$ . Suponiendo normas vectoriales y normas matriciales inducidas,  $J(x_0)$  no singular  $J \in \mathcal{L}ip_\gamma \mathcal{N}(x_0, r)$  y que existen constantes  $\beta, \eta > 0$  tal que

$$\|J(x_0)^{-1}F(x_0)\| \leq \eta, \quad \alpha = \beta\eta\gamma.$$

Entonces

i) Si  $\alpha \leq \frac{1}{2}$  y  $r > r_0 = \frac{(1-\sqrt{1-2\alpha})}{\beta\gamma}$ , entonces la sucesión generada por el métodos de Newton está bien definida y converge a  $x_*$  una única raíz de  $F(x)$  en  $\overline{\mathcal{N}(x_0, r_0)}$ .

ii) Si  $\alpha < \frac{1}{2}$  entonces  $x_*$  es la única raíz de  $F(x)$  en  $\mathcal{N}(x_0, r_1)$  donde  $r_1 = \min \left\{ r, \frac{(1+\sqrt{1-2\alpha})}{\beta\gamma} \right\}$   
y

$$\|x_k - x_*\| \leq \frac{\eta}{\alpha}(2\alpha)^k, \quad k \geq 0.$$



**Comentario:**

La principal diferencia entre el teorema de Kantorovich y el teorema de convergencia 2.1 es que en el teorema 2.2 no se tiene ningún tipo de información acerca de la solución  $x_*$  y la no singularidad de  $J(x_*)$ . El precio que se paga por esa falta de información es que sólo se obtiene convergencia  $r$ -cuadrática y no se tiene información de la manera que mejora cada iteración.

**Idea de la demostración:**

Siguiendo a Dennis-Schnabel [12], dejamos la demostración como un ejercicio, dando aquí una idea de los pasos que se deben seguir para completar la misma. Los estudiantes interesados en este tema también pueden consultar [17, 27] y [29].

Se utiliza la técnica de mayorización, construyendo una sucesión  $\{t_k\}$  de números reales que satisface las condiciones del teorema de Kantorovich.

- 1) Definir una función  $f : \mathbb{R} \rightarrow \mathbb{R}$  así

$$f(t) = \frac{\gamma}{2}t^2 - \frac{1}{\beta}t + \frac{\eta}{\beta}, \quad t_0 = 0.$$

Verificar que esta función satisface las condiciones del teorema 2.2:

- 1a)  $t_k \rightarrow t_* = \frac{1 - \sqrt{1 - 2\alpha}}{\beta\gamma}$ .  
 1b)  $|t_{k+1} - t_*| \leq 2^{k-1}\beta\gamma|t_k - t_*|^2$ .

- 2) Entonces probar que

- 2a)  $\|J(x_k)^{-1}\| \leq |f(t_k)^{-1}|$ .  
 2b)  $\|x_{k+1} - x_k\| \leq |t_{k+1} - t_k|$ .

- 3) Luego probar que  $\{x_k\}$  es una sucesión de Cauchy, que  $x_k \rightarrow x_*$  y que del hecho que  $\{t_n\}$  converge  $q$ -cuadráticamente, los errores están acotados para todo  $k$ .

**2.3.5 Teorema de punto fijo**

El método de Newton pertenece a una familia de métodos iterativos que satisfacen  $x_{k+1} = G(x_k)$ . El siguiente teorema establece condiciones sobre  $G$  bajo las cuales la sucesión  $\{x_k\}$  converge  $q$ -linealmente a  $x_*$  desde un punto  $x_0 \in \mathbb{D}$ . Además  $x_* \in \mathbb{D}$  es único tal que  $G(x_*) = x_*$ .

**Teorema 2.3** Sea  $G : \mathbb{D} \rightarrow \mathbb{D}$ ,  $\mathbb{D} \subset \mathbb{R}^n$  cerrado. Si para alguna norma existe  $\alpha \in [0, 1)$  tal que para todo  $x, y \in \mathbb{D}$

$$\|G(x) - G(y)\| \leq \alpha\|x - y\|,$$

entonces:

- a) Existe un único  $x_* \in \mathbb{D}$  tal que  $G(x_*) = x_*$ .
- b) Para todo  $x_0 \in \mathbb{D}$  la sucesión  $\{x_k\}$  generada por  $x_{k+1} = G(x_k)$ , permanece en  $\mathbb{D}$  y converge  $q$ -linealmente a  $x_*$  con constante  $\alpha$ .
- c) Para todo  $\eta$  tal que  $\|G(x_0) - x_0\| \leq \eta$  se cumple que  $\|x_k - x_*\| \leq \frac{\eta\alpha}{1-\alpha}$  para todo  $k = 0, 1, \dots$ .

### 2.3.6 Ejercicios

**Ejercicio 2.1** Muestre con el siguiente ejemplo que una sucesión puede converger linealmente en una norma y no en otra.

Sea  $\{x_k\} \in \mathbb{R}^2$ , definida como sigue

$$x_k = \begin{cases} (2^{-k}, 0)^T & \text{si } k \text{ es impar} \\ \sqrt{2}(2^{-k}, 2^{-k})^T & \text{si } k \text{ es par.} \end{cases}$$

Use  $\|\cdot\|_2$  y  $\|\cdot\|_\infty$ .

**Ejercicio 2.2** Sea el sistema

$$\begin{aligned} 10x + \sin(x + y) &= 1 \\ 8y - \cos^2(z - y) &= 1 \\ 12z + \sin z &= 1 \end{aligned}$$

- a) Comenzando de  $(\frac{1}{10}, \frac{1}{4}, \frac{1}{12})^T$ , siga una iteración del método de Newton.
- b) Sugiera un método explícito  $x_{k+1} = G(x_k)$ , donde no sea necesario invertir  $G$ , que resuelve el sistema no lineal.
- c) Comenzando con  $(\frac{1}{10}, \frac{1}{4}, \frac{1}{12})^T$ , estime cuantas iteraciones se necesitan para obtener una aproximación con seis lugares decimales correctos utilizando el método propuesto en b).

## Clase 3

# Métodos cuasi-Newton para SNL

Una de las desventajas del método de Newton para sistemas no lineales es el cálculo de la matriz Jacobiana, lo que requiere una cantidad considerable de evaluaciones de funciones. En esta clase veremos dos estrategias, para evitar el cálculo de las derivadas.

### 3.1 Método de Newton/diferencias finitas

En el caso de una variable  $f'(x)$  es aproximada por

$$a = \frac{f(x+h) - f(x)}{h}$$

donde  $h$  es una cantidad tal que  $|f'(x) - a| \approx \mathcal{O}(h)$ .

Para el caso  $n$ -dimensional es razonable aproximar la componente  $\frac{\partial f_i}{\partial x_j}$  por

$$a_{ij} = \frac{f_i(x + h\mathbf{e}_j) - f_i(x)}{h}, \quad i, j = 1, \dots, n$$

donde  $\mathbf{e}_j$  es el  $j$ -ésimo vector canónico. Esto es equivalente a aproximar la  $j$ -ésima columna de  $J(x)$  por el vector

$$A_{.j} = \frac{F(x + h\mathbf{e}_j) - F(x)}{h}, \quad j = 1, \dots, n.$$

Utilizando lema 2.1 se puede probar fácilmente

$$\begin{aligned} \|A_{.j} - J(x)_{.j}\| &\leq C_1|h| \\ \|A - J(x)\|_1 &\leq C_1|h|. \end{aligned}$$

Esto nos permite construir el siguiente algoritmo

**Algoritmo 3.1 (Método de Newton/diferencias finitas)**

Sea  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $x_0 \in \mathbb{R}^n$  un punto inicial,

**For**  $k = 0, \dots$ , hasta convergencia **do**

1. Para  $j = 1, \dots, n$  calcular las columnas de la aproximación del Jacobiano,

$$(A_k)_{.j} = \frac{F(x_k + h_k e_j) - F(x_k)}{h_k}.$$

2. Resolver el sistema lineal  $A_k s = -F(x_k)$ .
3. Definir  $x_{k+1} = x_k + s_k$ .

**end**

**Teorema 3.1 (Elección de  $h_k$ )**

Sean  $F, x_*$ ,  $x_0$  como en el teorema 2.1. Si  $\{h_k\}$  es una sucesión de números reales tal que  $0 < |h_k| \leq \eta$  para algún  $\eta > 0$ , la sucesión generada por el método de Newton con diferencias finitas hacia adelante converge a  $x_*$   $q$ -linealmente. Además

1. Si  $h_k \rightarrow 0$  la convergencia es  $q$ -superlineal.
2. Si existe  $c_1 > 0$  tal que  $|h_k| \leq c_1 \|x_k - x_*\|$  o equivalentemente si existe  $c_2 > 0$  tal que  $|h_k| \leq c_2 \|F(x_k)\|$  la convergencia es  $q$ -cuadrática.

**Demostración:** Ver [12]. ■

Como en el caso de una variable este teorema no indica exactamente como elegir el tamaño del incremento  $h_k$  en práctica.

Es claro que hay dos fuentes de error en conflicto.

- Si  $A_{.j}$  es una aproximación de  $J(x)_{.j}$ , el error es  $\mathcal{O}(|h|)$ , esto sugiere tomar  $h$ , pequeño.
- Sin embargo otro error que se introduce en el cálculo de  $A_{.j}$  es causado al evaluar el numerador  $F(x + h e_j) - F(x)$ . Este error, digamos  $\delta$  es cada vez peor para valores pequeños de  $h$ . Al dividir por  $h$  el error en el cociente es  $\frac{\delta}{h}$ . El error  $\delta$  que resulta de la inexactitud de los valores de  $F$  y de errores de cancelación, se supone que es una fracción de  $\|F(x)\|$ .

Entonces  $h$  debe elegirse de modo que exista un balance entre  $\mathcal{O}(|h|)$  y  $\mathcal{O}\left(\frac{\|F(x)\|}{|h|}\right)$  en la aproximación  $A_j$  de la  $j$ -ésima columna de  $J(x)$ .

Es razonable, que si  $F(x)$  tiene  $t$ -dígitos confiables, entonces

$$F(x + h\mathbf{e}_j) - F(x)$$

debería diferir de  $F(x)$  en la última mitad de estos dígitos. Más precisamente si el error relativo en el computo de  $F(x)$  es  $\eta$ , entonces

$$\frac{\|F(x + h\mathbf{e}_j) - F(x)\|}{\|F(x)\|} \leq \sqrt{\eta}, \quad j = 1, \dots, n.$$

En ausencia de una mejor información para conseguir esto, es razonable perturbar cada componente  $x_j$  por su propio  $h_j$ . Usualmente se toma  $h_j = \sqrt{\varepsilon_M}|x_j|$ , para todo  $j$ .

**Ejemplo 3.1** Ver [12].

Consideremos

$$f(x_1, x_2) = \begin{pmatrix} x_1^2 + x_2^2 - 2 \\ e^{x_1-1} + x_2^3 - 2 \end{pmatrix}.$$

Una solución es  $x_\star = (1, 1)^T$ . Aplicamos el algoritmo 3.1 comenzando desde  $x_0 = (2, 3)$  y  $h_j = 10^{-7}|x_j|$ ,  $j = 1, 2$ .

$$J_0 = \begin{pmatrix} 4.0000003309615 & 6.0000002122252 \\ 2.7182824169358 & 27.000002470838 \end{pmatrix}.$$

$$\begin{array}{ll} x_1 = \begin{pmatrix} 0.57465515450268 \\ 2.11689666735234 \end{pmatrix} & \|x_1 - x_\star\|_\infty \approx 10 \\ x_2 = \begin{pmatrix} 0.31178738552306 \\ 1.52419810116335 \end{pmatrix} & \|x_2 - x_\star\|_\infty \approx 7 \times 10^{-1} \\ x_3 = \begin{pmatrix} 1.484138151178 \\ 1.1464781318492 \end{pmatrix} & \|x_3 - x_\star\|_\infty \approx 5 \times 10^{-1} \\ x_4 = \begin{pmatrix} 1.0592958450507 \\ 1.03481950922365 \end{pmatrix} & \|x_4 - x_\star\|_\infty \approx 6 \times 10^{-2} \\ x_5 = \begin{pmatrix} 1.0008031056081 \\ 1.0014625533494 \end{pmatrix} & \|x_5 - x_\star\|_\infty \approx 2 \times 10^{-3} \\ x_6 = \begin{pmatrix} 0.99999872173640 \\ 1.0000026674316 \end{pmatrix} & \|x_6 - x_\star\|_\infty \approx 2 \times 10^{-6} \\ x_7 = \begin{pmatrix} 0.99999999999535 \\ 1.00000000000091 \end{pmatrix} & \|x_7 - x_\star\|_\infty \approx 9 \times 10^{-12}. \end{array}$$

### 3.2 Métodos secantes para sistemas no lineales

Recordemos que el método de Newton aplicado a  $F(x) = 0$ , en cada iteración resuelve un sistema lineal

$$F'(x_c)s = -F(x_c),$$

si la solución de este sistema es  $s_c$ , el nuevo iterado es

$$x_+ = x_c + s_c.$$

La iteración de Newton proviene de aproximar  $F(x)$  por el modelo lineal (2.4) alrededor del iterado actual  $x_c$ , esto es

$$M_c(x) = F(x_c) + F'(x_c)(x - x_c).$$

Una desventaja del método de Newton es el cálculo de  $F'(x_c)$ . Por otra parte, si la matriz Jacobiana es aproximada utilizando diferencias finitas, nos encontramos con el inconveniente de tener que efectuar  $n^2$  evaluaciones de funciones.

Entonces, ¿es posible aproximar  $F'(x_c)$  o su inversa, cuando éstas no están disponibles, sin efectuar evaluaciones de funciones?

Recordemos que en el método de la secante para el caso unidimensional, considerábamos los *dos últimos iterados*,  $x_c$ ,  $x_+$ , y el modelo afín que aproxima  $f$  en un entorno de  $x_+$ ,

$$\begin{aligned} m_+(x) &= f(x_+) + a_+(x - x_+) \\ a_+ &= \frac{f(x_+) - f(x_c)}{x_+ - x_c}. \end{aligned}$$

Observemos:

$$m_+(x_+) = f(x_+) \tag{3.1}$$

$$m_+(x_c) = f(x_c) \tag{3.2}$$

$$m_+(x_{++}) = 0. \tag{3.3}$$

El precio que se paga por usar esta aproximación es que se pierde la rápida convergencia cuadrática, obteniéndose sólo convergencia de orden  $p = \frac{1+\sqrt{5}}{2}$ .

Para el caso  $n > 1$ , se procede en forma similar. Consideremos  $x_c, x_+ \in \mathbb{R}^n$ . El modelo lineal

$$M_+(x) = F(x_+) + A_+(x - x_+),$$

satisface  $M_+(x_+) = F(x_+)$  y EXIGIMOS QUE  $A_+$  SEA TAL QUE

$$M_+(x_c) = F(x_c) \quad (3.4)$$

entonces

$$\begin{aligned} M_+(x_c) &= F(x_+) + A_+(x_c - x_+) \\ F(x_c) &= F(x_+) + A_+(x_c - x_+) \\ A_+(x_+ - x_c) &= F(x_+) - F(x_c). \end{aligned}$$

Dado que  $x_+ - x_c = s_c$  es el paso y definiendo  $y_c = F(x_+) - F(x_c)$  se tiene la **ecuación de la secante**

$$A_+ s_c = y_c. \quad (3.5)$$

Este es un sistema de  $n$  ecuaciones y  $n^2$  incógnitas, las entradas de la matriz  $A_+ \in \mathbb{R}^{n \times n}$ . Por lo tanto no se tiene una solución única.

¿Cómo hallar una matriz  $A_+$ ?

Dado que en la iteración actual no se tiene ninguna información acerca de la matriz Jacobiana o del modelo, hay que conservar la información que se tiene de las iteraciones previas. Por lo tanto el objetivo es:

*ELEGIR  $A_+$  TRATANDO DE MINIMIZAR EL CAMBIO EN EL MODELO LINEAL SATISFACIENDO LA ECUACION DE LA SECANTE.*

**Teorema 3.2** Sean  $x_c, x_+ \in \mathbb{R}^n$  tal que  $A_+ s_c = y_c$ , entonces

$$M_+(x) - M_c(x) = (A_+ - A_c)(x - x_c)$$

para todo  $x \in \mathcal{E}(x_+)$ .

**Demostración:**

$$\begin{aligned}
 M_+(x) - M_c(x) &= F(x_+) + A_+(x - x_+) - (F(x_c) + A_c(x - x_c)) \\
 &= F(x_+) - F(x_c) + A_+x - A_+x_+ - A_cx + A_cx_c + A_+x_c \\
 &\quad - A_+x_c \\
 &= F(x_+) - F(x_c) - A_+(x_+ - x_c) + A_+(x - x_c) - A_c(x - x_c) \\
 &= y_c - A_+s_c + (A_+ - A_c)(x - x_c) \\
 &= (A_+ - A_c)(x - x_c).
 \end{aligned}$$

Ahora bien, dado  $x \in \mathbb{R}^n$ ,  $x - x_c$  se puede descomponer de manera única como

$$x - x_c = \alpha s_c + \mathbf{n}, \quad \alpha \neq 0,$$

con  $s_c^T \mathbf{n} = 0$ . Entonces

$$\begin{aligned}
 M_+(x) - M_c(x) &= (A_+ - A_c)(x - x_c) \\
 &= (A_+ - A_c)(\alpha s_c + \mathbf{n}) \\
 &= \alpha(A_+ - A_c)s_c + (A_+ - A_c)\mathbf{n} \\
 &= \alpha(y_c - A_c s_c) + (A_+ - A_c)\mathbf{n}.
 \end{aligned} \tag{3.6}$$

En ecuación (3.6), el primer sumando involucra información conocida, luego para que el cambio en el modelo afín, al pasar de una iteración a otra, sea mínima en un entorno del iterado actual  $x_+$  se debe elegir  $A_+$  de modo que

$$\begin{cases} (A_+ - A_c)\mathbf{n} = 0 \\ s_c^T \mathbf{n} = 0. \end{cases}$$

Por lo tanto  $A_+$  no debe diferir de  $A_c$  en el complemento ortogonal del subespacio generado por la dirección del paso  $s_c$ . Esto significa que  $A_+ - A_c$  debe ser una matriz de rango 1. Luego existe un par de vectores  $u, v \in \mathbb{R}^n$  tal que

$$\begin{aligned}
 A_+ - A_c &= uv^T \\
 A_+ &= A_c + uv^T
 \end{aligned} \tag{3.7}$$

¿ Cuáles son los vectores  $u$  y  $v$  ?



Determinamos  $u$ . En (3.7), calculamos

$$\begin{aligned} A_+ s_c &= A_c s_c + (uv^T) s_c \\ &= A_c s_c + u(v^T s_c) \\ u &= \frac{y_c - A_c s_c}{v^T s_c}. \end{aligned} \quad (3.8)$$

**Conclusión:** Todas las actualizaciones o correcciones secantes de  $A_c$  están dadas por

$$A_+ = A_c + \frac{(y_c - A_c s_c) v^T}{v^T s_c} \quad (3.9)$$

para alguna elección de  $v$ .

### 3.2.1 Método de Broyden

Si en (3.9) se elige  $v$  como el paso  $s_c$ , se tiene la **actualización de Broyden**

$$A_+ = A_c + \frac{(y_c - A_c s_c) s_c^T}{s_c^T s_c} \quad (3.10)$$

Esta actualización permite definir el método de Broyden que es quizá después del método de Newton el método más popular para resolver sistemas no lineales de ecuaciones algebraicas.

¿Qué significa el término *actualizar* ?

Actualizar la matriz indica que no se está aproximando  $F'(x_+)$  ignorando  $F'(x_c)$ , sino que la aproximación  $A_c$  de  $F'(x_c)$  está siendo corregida de modo que  $A_+$  sea una aproximación de  $F'(x_+)$ . Establecemos el algoritmo :

**Algoritmo 3.2 (Método de Broyden)**

Sea  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $x_0 \in \mathbb{R}^n$  un punto inicial,  $A_0 \in \mathbb{R}^{n \times n}$  una matriz inicial,

**for**  $k = 0, \dots, \dots$ , hasta convergencia **repeat**

1. Resolver el sistema lineal  $A_k s = -F(x_k)$ .
2. Definir  $x_{k+1} = x_k + s_k$ .
3. Calcular  $y_k = F(x_{k+1}) - F(x_k)$ .
4. Actualizar la matriz

$$A_{k+1} = A_k + \frac{(y_k - A_k s_k) s_k^T}{s_k^T s_k}.$$

**end**

El método trabaja bien localmente como garantiza el siguiente teorema

**Teorema 3.3** (convergencia del método de Broyden)

Sea  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ . Bajo la hipótesis del teorema de convergencia del método de Newton, si  $x_0 \in \mathbb{R}^n$  es próximo a  $x_*$  (esto es  $\|x_0 - x_*\|_2 \leq \delta$  para algún  $\delta$ ) y  $A_0$  próxima a  $F'(x_*)$  la sucesión  $\{x_k\}$  generada por el método de Broyden converge a  $x_*$   $q$ -superlinealmente.

Que  $A_0$  sea próxima a  $F'(x_*)$  significa que en la norma  $\ell_2$ ,  $\|A_0 - F'(x_*)\|_2 \leq \eta$ , para algún  $\eta > 0$ .

Usualmente en práctica se elige  $A_0 = I$ ,  $A_0 = D$  (matriz diagonal) o bien  $A_0 = F'(x_0)$  ( ya sea analítica o calculada por diferencias finitas ).

**Ejemplo 3.2** Ver [12].

Consideremos

$$F(x_1, x_2) = \begin{pmatrix} x_1^2 + x_2^2 - 2 \\ e^{x_1-1} + x_2^3 - 2 \end{pmatrix}.$$

Una solución es  $x_* = (1, 1)^T$ . Aplicamos el algoritmo 3.2 comenzando desde  $x_0 = (1.5, 2)$  con  $A_0 = J(x_0)$ .

$$x_0 = \begin{pmatrix} 1.50000000 \\ 2.00000000 \end{pmatrix} \qquad \|x_0 - x_*\|_\infty = 1.$$

$$\begin{array}{ll}
x_1 = \begin{pmatrix} 0.8060692 \\ 1.457948 \end{pmatrix} & \|x_1 - x_\star\|_\infty \approx 5 \times 10^{-1} \\
x_2 = \begin{pmatrix} 0.7410741 \\ 1.277067 \end{pmatrix} & \|x_2 - x_\star\|_\infty \approx 3 \times 10^{-1} \\
x_3 = \begin{pmatrix} 1.8022786 \\ 1.159900 \end{pmatrix} & \|x_3 - x_\star\|_\infty \approx 9 \times 10^{-1} \\
x_4 = \begin{pmatrix} 1.9294701 \\ 1.070406 \end{pmatrix} & \|x_4 - x_\star\|_\infty \approx 9 \times 10^{-1} \\
x_5 = \begin{pmatrix} 1.004003 \\ 1.009609 \end{pmatrix} & \|x_5 - x_\star\|_\infty \approx 9 \times 10^{-3} \\
x_6 = \begin{pmatrix} 1.003084 \\ 0.9992213 \end{pmatrix} & \|x_6 - x_\star\|_\infty \approx 6 \times 10^{-3} \\
x_7 = \begin{pmatrix} 1.000543 \\ 0.9996855 \end{pmatrix} & \|x_7 - x_\star\|_\infty \approx 6 \times 10^{-4} \\
x_8 = \begin{pmatrix} 0.99999818 \\ 1.00000000389 \end{pmatrix} & \|x_8 - x_\star\|_\infty \approx 2 \times 10^{-6} \\
x_9 = \begin{pmatrix} 0.9999999885 \\ 0.999999999544 \end{pmatrix} & \|x_9 - x_\star\|_\infty \approx 2 \times 10^{-8} \\
x_{10} = \begin{pmatrix} 0.9999999999474 \\ 0.9999999999998 \end{pmatrix} & \|x_{10} - x_\star\|_\infty \approx 2 \times 10^{-11} \\
x_{11} = \begin{pmatrix} 1.0 \\ 1.0 \end{pmatrix} & \|x_{11} - x_\star\|_\infty = 0.
\end{array}$$

Se observa que la convergencia es  $q$ -superlineal. Notemos también que la aproximación final del Jacobiano generada por el método de Broyden es

$$A_{10} = \begin{pmatrix} 1.999137 & 2.021829 \\ 0.9995643 & 3.011004 \end{pmatrix},$$

mientras que

$$J(x_\star) = \begin{pmatrix} 2 & 2 \\ 1 & 3 \end{pmatrix}.$$

Se observa que  $A_{10}$  tiene un error relativo máximo del 1.1 % como aproximación de  $J(x_\star)$ . Este comportamiento es típico de la aproximación final del Jacobiano generado por el método de Broyden. Sin embargo es posible encontrar ejemplos donde la sucesión  $\{A_k\}$  no converge a  $J(x_\star)$ .

### Ventajas del método de Broyden

- 1) Rápida convergencia local si  $x_0$  y  $A_0$  son elegidos adecuadamente y  $J(x_*)$  es no singular, (si bien no tan rápida como la velocidad que se puede observar en el método de Newton).
- 2) No necesita el cálculo de derivadas.

### Desventajas del método de Broyden

- 1) El método no converge globalmente.
- 2) En cada iteración necesita resolver un sistema lineal, con los inconvenientes que esto puede provocar.

### 3.2.2 Ejercicios

**Ejercicio 3.1** Sea  $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ , definida como sigue

$$F(x_1, x_2) = \begin{pmatrix} x_1 + x_2 - 3 \\ x_1^2 + x_2^2 - 9 \end{pmatrix},$$

la cual tiene raíces  $(0, 3)^T$  y  $(3, 0)^T$ . Sea  $x_0 = (1, 5)^T$ .

- a) Halle una raíz utilizando el método de Newton.
- b) Repita a) utilizando el método de Broyden comenzado con  $x_0$  y  $A_0 = J(x_0)$ .
- c) Muestre que la sucesión  $\{A_k\}$  generada por el método de Broyden no converge a  $J(x_*)$ .

**Ejercicio 3.2** Considere el sistema no lineal

$$\begin{aligned} x^3 + y^3 &= 35 \\ 2y^3 + z^4 &= 69 \\ 3x^5 + 10z^2 &= 770. \end{aligned}$$

Halle la solución del sistema no lineal que es próxima a  $x = 3$ ,  $y = 3$ ,  $z = 2$ ,

- a) Utilizando el método de Newton.
- b) Utilizando el método de Broyden y  $J(x_0)$ , como matriz inicial.
- c) Utilizando el método de Broyden e  $I$ , como matriz inicial.
- d) Compare sus resultados.

# Clase 4

## Minimización sin restricciones

Consideremos el problema

$$\min_{x \in \mathbb{R}^n} f(x),$$

siendo  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ . Vamos a establecer condiciones necesarias y suficientes para que  $x_\star \in \mathbb{R}^n$  sea un mínimo local para  $f$ . Estas condiciones son la clave para los algoritmos que vamos a desarrollar.

### 4.1 Condiciones de optimalidad

#### **Teorema 4.1 (Condición necesaria de primer orden)**

Sea  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $f \in C^1(\mathbb{D})$ ,  $\mathbb{D} \subset \mathbb{R}^n$  abierto. Si  $x_\star \in \mathbb{D}$  es un mínimo local de  $f$  entonces  $\nabla f(x_\star) = 0$ .

**Demostración:** Sea  $d \in \mathbb{R}^n$  un vector arbitrario,  $d \neq 0$ . Sea

$$\phi : \mathbb{R} \rightarrow \mathbb{R}$$

definida como:

$$\phi(t) = f(x_\star + td).$$

Calculamos,

$$\phi'(t) = \nabla f(x_\star + td)^T d. \tag{4.1}$$

Como  $x_\star$  es mínimo local de  $f$ , entonces

$$\begin{aligned} f(x_\star) &< f(x_\star + td) \quad \forall d \neq 0 \\ \phi(0) &< \phi(t), \end{aligned}$$

de donde  $t_\star = 0$  es un mínimo local de  $\phi$ , y por lo tanto  $\phi'(0) = 0$ . En (4.1) resulta,

$$\phi'(0) = \nabla f(x_\star)^T d \quad (4.2)$$

Como  $\nabla f(x_\star) \not\perp d$  y  $d \neq 0$ , de (4.2) resulta que  $\nabla f(x_\star) = 0$ . ■

La siguiente, es otra demostración del lema que es interesante pues permite caracterizar una clase de algoritmos que resuelven el problema MSR: *los métodos de descenso*.

**Demostración:** (por el absurdo).

Supongamos que  $\nabla f(x_\star) \neq 0$ , entonces es posible elegir una dirección  $p \in \mathbb{R}^n$ ,  $p \neq 0$  tal que  $x + p \in \mathbb{D}$  y  $\nabla f(x_\star)^T p < 0$ .

Como  $f \in C^1(\mathbb{D})$ , entonces para todo  $t \in (0, \tau)$ , para algún  $\tau$  se tiene

$$\nabla f(x_\star + tp)^T p < 0.$$

Luego

$$f(x_\star + \tau p) - f(x_\star) = \int_0^\tau \nabla f(x_\star + tp)^T p dt < 0,$$

de donde  $x_\star$  no puede ser un mínimo para  $f$ . ■

Esta demostración, básicamente sugiere que si en un punto  $x_c$ ,  $\nabla f(x_c) \neq 0$ , se puede elegir una dirección  $p_c \neq 0$  tal que  $\nabla f(x_c)^T p_c < 0$  que permitirá hallar una nueva y mejor aproximación a  $x_\star$ .

#### **Teorema 4.2 (Condición necesaria de segundo orden)**

Sea  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $f \in C^2(\mathbb{D})$ ,  $\mathbb{D} \subset \mathbb{R}^n$  abierto y convexo. Si  $x_\star \in \mathbb{D}$  es un mínimo local de  $f$ , entonces

- a)  $\nabla f(x_\star) = 0$
- b)  $\nabla^2 f(x_\star)$  es semidefinida positiva.

**Demostración:** Consideremos nuevamente  $\phi(t) = f(x_\star + td)$  para todo  $d \neq 0$ . como  $x_\star \in \mathbb{D}$  es un mínimo local de  $f$ , el punto  $t_\star = 0$  es un mínimo local de  $\phi$ , entonces

$$\phi''(0) \geq 0 \quad (4.3)$$

$$\begin{aligned} \phi'(t) &= \nabla f(x_\star + td)^T d \\ \phi''(t) &= d^T \nabla^2 f(x_\star + td) d \\ \phi''(0) &= d^T \nabla^2 f(x_\star) d > 0 \end{aligned} \quad (4.4)$$

donde la desigualdad (4.4) se deduce a partir de (4.3). ■

**Teorema 4.3 (Condición suficiente de segundo orden)**

Sea  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $f \in C^2(\mathbb{D})$ ,  $\mathbb{D} \subset \mathbb{R}^n$  abierto y convexo. Si  $x_* \in \mathbb{D}$  tal que  $\nabla f(x_*) = 0$  y  $\nabla^2 f(x_*)$  es definida positiva, entonces  $x_*$  es un mínimo local estricto de  $f$ .

**Demostración:** Como  $\nabla^2 f(x_*)$  es definida positiva, debido a la continuidad de  $\nabla^2 f(x)$ , existe un entorno  $\mathcal{U}(x_*) \subset \mathbb{D}$  tal que  $\nabla^2 f(x)$  es definida positiva para todo  $x \in \mathcal{U}(x_*)$ . Entonces para todo  $p$  tal que  $x_* + p \in \mathbb{D}$ ,

$$\begin{aligned} f(x_* + p) &= f(x_*) + \nabla f(x_*)^T p + \frac{1}{2} p^T \nabla^2 f(z) p \\ &= f(x_*) + \frac{1}{2} p^T \nabla^2 f(z) p \\ &> f(x_*), \end{aligned}$$

con  $z = \alpha x_* + (1 - \alpha)(x_* + p)$ ,  $1 \leq \alpha \leq 1$ . En consecuencia  $x_*$  es un mínimo local estricto de  $f$ . ■

**Corolario 4.1** Sea  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $f \in C^2(\mathbb{D})$ ,  $\mathbb{D} \subset \mathbb{R}^n$  abierto y convexo. Si  $x_* \in \mathbb{D}$  tal que  $\nabla f(x_*) = 0$  y  $\nabla^2 f(x_*) \in \mathcal{L}ip_L(\mathcal{U}(x_*))$  con  $\nabla^2 f(x_*)$  no singular, entonces  $x_*$  es un mínimo local estricto de  $f$  si y sólo si  $\nabla^2 f(x_*)$  es definida positiva.

**Demostración:**

$\Leftarrow$ ) obvio.

$\Rightarrow$ ) Si  $x_*$  es un mínimo local de  $f$ , entonces  $\nabla f(x_*) = 0$  y  $\nabla^2 f(x_*)$  es semidefinida positiva. Como  $\nabla^2 f(x_*)$  es no singular, entonces  $\nabla^2 f(x_*)$  es definida positiva. ■

**Comentario:** Es claro que las condiciones necesarias y suficientes para que un punto  $x_*$  sea un máximo local de  $f$  son simplemente que  $x_*$  sea un mínimo local de  $-f$ .

## 4.2 Direcciones de descenso

**Definición 4.1 (Dirección de descenso)**

Sea  $f \in C^1$ , el vector  $p \in \mathbb{R}^n$ ,  $p \neq 0$  se dice que es una dirección de descenso para  $f$  desde  $\bar{x}$  si existe  $\tau > 0$  tal que para todo  $t \in (0, \tau)$ ,

$$f(\bar{x} + tp) < f(\bar{x}).$$

**Proposición 4.1** Si  $\nabla f(\bar{x})^T p < 0$  entonces  $p$  es una dirección de descenso para  $f$  en  $x = \bar{x}$ .

La recíproca no es cierta.

**Demostración:** Si  $\nabla f(\bar{x})^T p < 0$ , entonces para todo  $t \in (0, \tau)$ ,

$$\frac{f(\bar{x} + tp) - f(\bar{x})}{t} < 0$$

de donde

$$f(\bar{x} + tp) < f(\bar{x}).$$

Con el siguiente ejemplo mostramos que la recíproca no es cierta.

**Ejemplo 4.1** Sean

$$f(x_1, x_2) = -x_1^2 - x_2^2, \quad \bar{x} = (0, 0)^T, \quad p = (1, 0)^T \neq (0, 0)^T.$$

- $p$  es una dirección de descenso.

$$\begin{aligned} f(\bar{x}) &= f(0, 0) = 0 \\ f(\bar{x} + tp) &= f(t, 0) = -t^2 < 0 \end{aligned}$$

para todo  $t \in (0, \tau)$ , para algún  $\tau > 0$ . Entonces

$$f(\bar{x} + tp) < f(\bar{x}).$$

- Claramente  $\nabla f(0, 0) = (0, 0)^T$  y por lo tanto

$$\nabla f(\bar{x})^T p \not< 0.$$

■

Establecido el concepto de dirección de descenso y con ayuda de la segunda demostración del teorema 4.1, nos proponemos desarrollar algoritmos que a partir de un punto inicial dado, generen sucesiones que converjan a una solución del problema MRS.

Una iteración típica de un algoritmo de descenso es

**Algoritmo 4.1 (Iteración de cualquier método de descenso)**

Dada  $x_c \in \mathbb{R}^n$ , determinar  $p_c$  tal que  $\nabla f(x_c)^T p_c < 0$ .

- Definir el paso de prueba:  $s_c = \alpha_c p_c$ .
- Determinar si el paso es aceptable:  $x_+ = x_c + s_c$  con

$$f(x_c + s_c) \ll f(x_c)$$



### 4.2.1 Método de Newton

Dado  $x_c$ , alrededor de  $x_c$  la función  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  se puede aproximar por un modelo cuadrático

$$f(x) = f(x_c + s) \approx f(x_c) + \underbrace{\nabla f(x_c)^T s + \frac{1}{2} s^T \nabla^2 f(x_c) s}_{\equiv q_c(s)}.$$

Para el modelo  $q_c(s)$  una condición necesaria para que  $s$  sea un mínimo es  $\nabla q_c(s) = 0$ . Esto es

$$\nabla f(x_c) + \nabla^2 f(x_c) s = 0,$$

que es el modelo lineal de  $F(x) = \nabla f(x_c)$  alrededor de  $x_c$ . El paso de Newton es

$$s_c^N = -\nabla^2 f(x_c)^{-1} \nabla f(x_c). \tag{4.5}$$

A partir de esto, podemos construir el siguiente algoritmo

**Algoritmo 4.2 (Método de Newton para MSR)**  
 Sea  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $x_0 \in \mathbb{R}^n$  un punto inicial,  
**For**  $k = 0, \dots, \text{hasta convergencia}$  **do**

- Resolver el sistema lineal  $\nabla^2 f(x_k) s = -\nabla f(x_k)$ .
- Definir  $x_{k+1} = x_k + s_k$ .

**end**

#### Ventajas y desventajas del método

- Es claro que las ventajas y desventajas del método de Newton para el problema sin restricciones, son las mismas que para sistemas no lineales de ecuaciones algebraicas.
- si  $x_0$  es próximo al mínimo local de  $f$ , con  $\nabla^2 f(x_*)$  simétrica y definida positiva, la sucesión generada por el método converge a  $x_*$   $q$ -cuadráticamente.
- Si  $f$  es una cuadrática estrictamente convexa, entonces  $\nabla f(x)$  es afín y la solución se halla en una iteración.

- Además de las desventajas que el algoritmo comparte con el método de Newton para sistemas no lineales de ecuaciones algebraicas, se presenta un nuevo inconveniente. Esto es a las dificultades que el método no es globalmente convergente, que en cada iteración ahora se requiere no sólo las derivadas exactas para formar el gradiente, sino que también se requiere la matriz Hessiana, se agrega el problema que la aproximación que se obtiene, más que un mínimo para  $f$ , es un punto estacionario de  $f$ . De hecho en cada iteración, el algoritmo permite calcular  $x_{k+1}$  del cual no se puede decir si es un mínimo, máximo o punto de ensilladura del modelo.

Aún si  $x_c$  es un punto próximo a un mínimo de la función de modo que  $\nabla^2 f(x_c)$  sea definida positiva, es interesante *modificar* el algoritmo de modo que se adapte cuando  $\nabla^2 f(x_c)$  no es definida positiva. Existen dos modificaciones razonables del método de Newton para el caso en que  $\nabla^2 f(x_c)$  no sea definida positiva.

1. *Tratar de utilizar la forma de la función, (la curvatura)*. Esto significa explotar las direcciones de descenso de curvatura negativa, esto es aquellas direcciones  $p$  tal que

$$\begin{aligned} p^T \nabla^2 f(x_c) p &< 0 \\ \nabla f(x_c)^T p &< 0. \end{aligned}$$

2. *Cambiar el modelo* como sigue

$$q_c(s) = f(x_c) + \nabla f(x_c)^T s + \frac{1}{2} s^T [\nabla^2 f(x_c) + \mu_c I] s$$

donde  $\mu_c$  es elegido para que  $\nabla^2 f(x_c) + \mu_c I$  sea definida positiva y razonablemente bien condicionada.

**Algoritmo 4.3 (Método de Newton modificado)**

Sea  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $x_0 \in \mathbb{R}^n$  un punto inicial,

**For**  $k = 0, \dots, \text{hasta convergencia}$  **do**

- Definir  $H_k = \nabla^2 f(x_k) + \mu_k I$  donde

$$\mu_k \begin{cases} = 0 & \text{si } \nabla^2 f(x_k) \text{ es d.p.} \\ > 0 & \text{en otro caso.} \end{cases}$$

- Resolver el sistema lineal  $H_k s = -\nabla f(x_k)$ .
- Definir  $x_{k+1} = x_k + s_k$ .

**end**

Tanto en este algoritmo como en algoritmo 4.2, el sistema  $H_k s = -\nabla f(x_k)$  se resuelve efectuando una factorización de Cholesky.

$$L_k L_k^T s = -\nabla f(x_k),$$

donde  $L_k$  es triangular inferior.

### 3. Elección del parámetro $\mu_c$ :

Si se dá a  $\mu_c$  un valor grande, nos aseguramos que  $H_c$  es definida positiva, pero posiblemente resulte una matriz mal condicionada. Luego  $\mu_c$  no tiene que ser tan grande.

Si  $\nabla^2 f(x_c)$  no es definida positiva, tiene un autovalor real negativo. Entonces se elige  $\mu_c > |\lambda_{\min}|$ , siendo  $\lambda_{\min}$  el autovalor negativo de mayor magnitud de  $\nabla^2 f(x_c)$ .

4. Es claro que si  $x_0$  es elegido próximo a  $x_*$  con  $\nabla^2 f(x_*)$  definida positiva, se puede aplicar el teorema 2.1 y se tiene que la sucesión generada por el método de Newton converge  $q$ -cuadráticamente a una solución local del problema. Para el problema MSR las hipótesis estandar son

(N1) Existe  $x_* \in \mathbb{R}^n$  tal que  $\nabla f(x_*) = 0$ .

(N2)  $\nabla^2 f(x_*) \in \mathcal{Lip}_L(\mathcal{U}(x_*))$ .

(N3)  $\nabla^2 f(x_*)$  es definida positiva.

**Proposición 4.2** *sea  $H_c$  la matriz dada por el algoritmo 4.3. Entonces la dirección  $p_c = -H_c^{-1}\nabla f(x_c)$  es una dirección de descenso.*

**Demostración:** Ejercicio ■

**Comentarios:** Para terminar,

1. Se puede evitar el cálculo de las derivadas parciales para definir la matriz Hessiana, usando diferencias finitas. Los alumnos interesados son remitidos al texto de Dennis y Schnabel [12].
2. Los métodos secantes necesitan un tratamiento especial para el problema de minimización sin restricciones. La bibliografía sugerida incluye los siguientes textos [12, 2, 3, 14, 20] y las contribuciones [10, 11].
3. Para los problema de gran porte los alumnos interesados pueden leer [5, 8, 7].

## Clase 5

# Minimización con restricciones

Comencemos el tema de optimización con restricciones con un ejemplo.

### 5.1 Un ejemplo

Consideremos el problema de estimar la solución de una ecuación diferencial, de la cual se conoce que ajusta un conjunto de observaciones con ruidos  $\{(t_1, y_1), \dots, (t_m, y_m)\}$ . Supongamos, a modo de ejemplo, la siguiente ecuación diferencial

$$-y'' + e^y = F(t), \quad 0 \leq t \leq 2\pi. \quad (5.1)$$

Por simplicidad supongamos que  $F(t) = \sin t + e^{\sin t}$ , ya que  $\sin t$  es una solución particular de (5.1).

Problemas de este tipo, aunque mucho más complicado, aparecen en el área de meteorología cuando se pretenden hacer pronósticos acerca del clima de una región.

Hay que determinar la función objetivo y la región de factibilidad. Para determinar el conjunto de factibilidad  $\Omega$ , resolvemos la ecuación diferencial por medio de un método numérico. En este caso usamos diferencias finitas. Dividimos el intervalo  $[0, 2\pi]$  en  $N$  subintervalos del mismo tamaño, haciendo  $h = \frac{2\pi}{N}$  y discretizamos (5.1). En cada punto de la grilla  $t_i = ih$ ,  $i = 0, \dots, N$  se denota  $y(ih) = x_{i+1}$ , y por lo tanto

$$x = (x_1, x_2, \dots, x_{N+1})^T = (y(0), y(h), \dots, y(Nh))^T.$$

Entonces la discretización de la ecuación diferencial es

$$\frac{-x_{i-1} + 2x_i - x_{i+1}}{h^2} + e^{x_i} = F((i-1)h), \quad i = 2, \dots, N. \quad (5.2)$$

Como a priori se conoce que la solución satisface  $-1 \leq y(t) \leq 1$  se tiene que

$$-1 \leq x_i \leq 1, \quad i = 1, \dots, N+1. \quad (5.3)$$

Asumiendo que las abscisas observadas son puntos de la grilla,

$$t_1 = j_1 h, t_2 = j_2 h, \dots, t_m = j_m h,$$

la función objetivo es

$$f(x) = f(y(0), y(h), \dots, y(Nh)) = \sum_{l=1}^m \phi(|y(j_l h) - y_l|), \quad (5.4)$$

donde  $\phi$  es una medida de la desviación de la verdadera solución con respecto al valor observado.

Usualmente se usa  $\phi(z) = z^2$ . Sin embargo, si algunas observaciones están sujetas a errores muy grandes, se necesita una función de desvío que no este sujeta a valores que difieran tanto entre si. En [15] Gomes, Maciel y Martínez proponen usar para este ejemplo la siguiente función

$$\phi(t) = \begin{cases} z^2 & z \leq 0.25 \\ 0.5(z - 0.25) + 0.0625 & \text{en otro caso.} \end{cases}$$

Entonces el problema consiste en hallar  $x \in \mathbb{R}^{N+1}$  tal que

$$\begin{array}{ll} \text{Minimice} & f(x) = \sum_{l=1}^m \phi(|y(j_l h) - y_l|) \\ & x \in \Omega \end{array}$$

con

$$\Omega = \{x \in \mathbb{R}^{N+1} : x \text{ satisface (5.1), (5.2)}\}.$$

La función objetivo no tiene derivada segunda en aquellos puntos donde el argumento de  $\phi$  es 0.25.

## 5.2 Multiplicadores de Lagrange

El conjunto de restricciones del problema general de programación no lineal usualmente se especifica en términos de igualdades y desigualdades. Teniendo en cuenta esta estructura se pueden obtener toda una familia de condiciones de optimalidad involucrando variables auxiliares que son precisamente los **multiplicadores de Lagrange**. Estos multiplicadores no sólo son útiles para caracterizar las soluciones del problema, sino que también brindan información para analizar cuán sensible son la solución y la funcional frente a pequeñas variaciones en los datos. También pueden ser usados para deducir métodos computacionales para hallar una solución local del problema. Finalmente, ellos pueden ser vistos como variables de problemas auxiliares llamados *problemas duales*.

La teoría de los multiplicadores de Lagrange se puede desarrollar en varias formas. Aquí seguimos las notas de Tapia [36] y el libro de Fiacco y McCormick [13], que es un clásico en el área. Se debe mencionar que Bertsekas [3], desarrolla la teoría de los multiplicadores siguiendo dos puntos de vista diferentes, el de penalización y el de direcciones factibles. Tratando que la teoría sea tan general como sea posible, asumimos un espacio de Hilbert  $H$  no necesariamente de dimensión finita. Por lo tanto consideramos el problema

$$(PNL) \equiv \begin{cases} \text{Minimice} & f(x) \\ \text{s.a} & c_i(x) = 0, \quad i \in I \\ & c_j(x) \geq 0, \quad j \in D. \end{cases}$$

donde  $f, c_i, c_j : H \rightarrow \mathbb{R}$ ,  $I = \{1, \dots, m\}$ ,  $D = \{m + 1, \dots, m + p\}$ . Llamemos  $S$  a la región de factibilidad,

$$S = \{x \in H : c_i(x) = 0, i \in I, c_j(x) \geq 0, j \in D\}.$$

### 5.2.1 Geometría del problema

Recordemos que para el caso de minimización sin restricciones cualquier dirección  $d$ , tal que el ángulo que forma  $d$  con la dirección negativa del gradiente es menor que 90 es una dirección de descenso para  $f$  en  $x_c$ . Esto es

$$\nabla f(x_c)^T d < 0.$$

¿Cómo se extiende esta propiedad al caso con restricciones ?. Es claro que se debe tener en cuenta

### OPTIMALIDAD & FACTIBILIDAD

Es decir las direcciones no sólo deben ser de descenso sino que ellas deben permanecer factibles. Supongamos el caso  $I = \emptyset$ ,

$$S = \{x \in H : c_j(x) \geq 0, j \in D\}.$$

Para mantener factibilidad, notemos que para todo  $x$  tal que  $c_j(x) = 0, j \in D$ , los gradientes  $\nabla c_j(x)$  deben apuntar al interior de  $S$ .

Para fijar ideas comencemos suponiendo una restricción,  $C(x) \geq 0$ . Si  $x_c$  no es optimal, siempre es posible hallar una dirección factible  $d$ , que reduce el valor de  $f$ ,

$$\begin{aligned}\nabla f(x_c)^T d &< 0 \\ x_c + td &\in S.\end{aligned}$$

Si  $x_c$  es una solución del problema de optimización (esto es  $x_c$  es optimal), no existe ninguna dirección que reduce el valor de  $f$  y mantiene factibilidad. Observemos que  $\nabla f(x_c) = \mu \nabla C(x_c)$ ,  $\mu \geq 0$ .

Si se tienen más restricciones y  $x_c$  es optimal,  $\nabla f(x_c)$  es una combinación lineal positiva de  $\nabla c_j(x_c)$ ,  $j \in D$ ,

$$\nabla f(x_c) = \mu_1 \nabla c_1(x_c) + \mu_2 \nabla c_2(x_c) + \mu_3 \nabla c_3(x_c), \quad \mu_j \geq 0,$$

$$\nabla f(x_c) - \mu_1 \nabla c_1(x_c) - \mu_2 \nabla c_2(x_c) - \mu_3 \nabla c_3(x_c) = 0, \quad \mu_j \geq 0.$$

Notemos que el primer miembro es el gradiente de la función

$$f(x) - \mu_1 c_1(x) - \mu_2 c_2(x) - \mu_3 c_3(x),$$

em  $x = x_c$ .

## 5.2.2 Existencia de los multiplicadores de Lagrange

Consideremos el problema PNL.

### Definición 5.1 (Lagrangiano)

Asociada al problema PNL, existe una función

$$\ell : H \times \mathbb{R}^m \times \mathbb{R}^p \rightarrow \mathbb{R}, \quad \ell(x, \lambda, \mu) = f(x) + \sum_{i \in I} \lambda_i c_i(x) - \sum_{j \in D} \mu_j c_j(x).$$

La función  $\ell$  se dice Lagrangiano o función de Lagrange asociada al problema PNL. Las componentes de los vectores  $\lambda \in \mathbb{R}^m$ ,  $\mu \in \mathbb{R}^p$  se llaman multiplicadores de Lagrange o variables duales.

### Definición 5.2 (Restricciones activas)

Sea  $x_*$  un punto factible de PNL. Las restricciones  $c_j(x_*) = 0$ ,  $j \in D$  se dicen restricciones activas en  $x_*$ .

Denotamos el conjunto de restricciones activas con  $B(x_*)$ , esto es

$$B(x_*) = \{j : j \in D, c_j(x_*) = 0\}.$$



El lema de Farkas juega un rol crucial para probar la existencia de los multiplicadores de Lagrange.

**Lema 5.1 (Lema de Farkas)**

Sea  $H$  un espacio de Hilbert.  $a_0, a_1, \dots, a_q \in H$ . Las siguientes proposiciones son equivalentes:

$$(i) \quad \forall z \in H \langle z, a_i \rangle_H \geq 0, \quad i = 1, \dots, q \Rightarrow \langle z, a_0 \rangle \geq 0.$$

$$(ii) \quad \exists t_i \geq 0 \text{ tal que } a_0 = \sum_{i=1}^q t_i a_i.$$

Para el problema PNL consideremos  $x$  un punto factible y supongamos que  $f$  y  $c_i \in C^1(S)$ . Definimos los siguientes conjuntos.

$$Z_1(x) = \{z \in H : \langle z, \nabla c_i(x) \rangle_H = 0, \quad i \in I, \quad \langle z, \nabla c_j(x) \rangle_H \geq 0, \quad j \in B(x), \\ \langle z, \nabla f(x) \rangle_H \geq 0\}. \quad (5.5)$$

$$Z_2(x) = \{z \in H : \langle z, \nabla c_i(x) \rangle_H = 0, \quad i \in I, \quad \langle z, \nabla c_j(x) \rangle_H \geq 0, \quad j \in B(x), \\ \langle z, \nabla f(x) \rangle_H < 0\}. \quad (5.6)$$

**Lema 5.2 (Existencia de los multiplicadores de Lagrange)** Sea  $x_*$  un punto factible de PNL. Las siguientes condiciones son equivalentes:

$$(i) \quad Z_2(x_*) = \emptyset.$$

(ii) Existe  $\lambda_* \in \mathbb{R}^m$ ,  $\mu_* \in \mathbb{R}^p$  tal que las siguientes condiciones se satisfacen

**Condiciones de Karush-Kuhn-Tucker (KKT)**

a)  $\nabla_x \ell(x_*, \lambda_*, \mu_*) = 0$

b) Factibilidad:  $c_i(x_*) = 0, \quad i \in I; \quad c_j(x_*) \geq 0, \quad j \in D$

c) Complementaridad:  $(\mu_*)_j c_j(x_*) = 0, \quad j \in D$

d) No negatividad:  $(\mu_*)_j \geq 0, \quad j \in D.$

**Comentario:** Las condiciones KKT son un tema central en la teoría de PNL. Fueron deducidas en forma independiente por Kuhn y Tucker a fines de la década del 40 y por Karush en 1939. Desafortunadamente el trabajo de Karush fue ignorado mucho

tiempo. Aquellos interesados en el área de optimización el artículo de Kuhn [18] hace una revisión histórica de los orígenes de la programación no lineal.

**Demostración:**  $\Leftarrow$ ) Supongamos que las condiciones KKT se satisfacen, y que  $Z_2(x_*) \neq \emptyset$ . Entonces existe  $z \in H$  tal que  $\langle z, \nabla f(x_*) \rangle_H < 0$ . Calculamos

$$\begin{aligned} \langle z, \nabla f(x_*) \rangle_H &= \langle z, \nabla_x \ell(x_*, \lambda_*, \mu_*) - \sum_{i \in I} (\lambda_*)_i \nabla c_i(x_*) + \sum_{i \in D} (\mu_*)_i \nabla c_i(x_*) \rangle_H \\ &= - \sum_{i \in I} (\lambda_*)_i \langle z, \nabla c_i(x_*) \rangle_H + \sum_{i \in D} (\mu_*)_i \langle z, \nabla c_i(x_*) \rangle_H \\ &= \sum_{i \in D} (\mu_*)_i \langle z, \nabla c_i(x_*) \rangle_H, \end{aligned}$$

pues se cumple a) y  $z \in Z_2(x_*)$ . Entonces

$$0 > \langle z, \nabla f(x_*) \rangle_H = \sum_{i \in B(x_*)} (\mu_*)_i \langle z, \nabla c_i(x_*) \rangle_H + \sum_{i \notin B(x_*)} (\mu_*)_i \langle z, \nabla c_i(x_*) \rangle_H. \quad (5.7)$$

Si  $i \in B(x_*)$ , se tiene que  $c_i(x_*) = 0$  y por lo tanto  $\mu_i \geq 0$  y

$$\sum_{i \in B(x_*)} (\mu_*)_i \langle z, \nabla c_i(x_*) \rangle_H \geq 0.$$

Si  $i \notin B(x_*)$ , se tiene que  $c_i(x_*) > 0$  y por lo tanto  $\mu_i = 0$  y

$$\sum_{i \notin B(x_*)} (\mu_*)_i \langle z, \nabla c_i(x_*) \rangle_H \geq 0.$$

En cualquier caso en (5.7) resulta que

$$0 > \langle z, \nabla f(x_*) \rangle_H \geq 0,$$

que es un absurdo.

$\Rightarrow$ ) Para deducir la existencia de los multiplicadores de Lagrange utilizamos el lema de Farkas.

Suponemos  $Z_2(x_*) = \emptyset$ . Entonces si  $\langle z, \nabla c_i(x_*) \rangle_H = 0$ ,  $i \in I$ ,  $\langle z, \nabla c_j(x_*) \rangle_H \geq 0$ ,  $j \in B(x_*)$  se satisfacen, debe ser  $\langle z, \nabla f(x_*) \rangle_H \geq 0$ .

Observemos que estas desigualdades se pueden escribir así

$$\begin{aligned} \langle z, \nabla c_i(x_*) \rangle_H &\geq 0, & i \in I \\ \langle z, -\nabla c_i(x_*) \rangle_H &\geq 0, & i \in I \\ \langle z, \nabla c_j(x_*) \rangle_H &\geq 0, & j \in B(x_*) \end{aligned}$$

entonces  $\langle z, \nabla f(x_*) \rangle_H \geq 0$ .

El lema de Farkas permite asegurar que existen  $\mu_i \geq 0, t_i \geq 0, t'_i \geq 0$  tal que

$$\begin{aligned} \nabla f(x_*) &= \sum_{i \in B(x_*)} \mu_i \nabla c_i(x_*) + \sum_{i \in I} t_i \nabla c_i(x_*) - \sum_{i \in I} t'_i \nabla c_i(x_*) \\ &= \sum_{i \in B(x_*)} \mu_i \nabla c_i(x_*) + \sum_{i \in I} (t_i - t'_i) \nabla c_i(x_*). \end{aligned} \quad (5.8)$$

Definiendo

$$\begin{aligned} \mu_* &= \begin{pmatrix} \mu_{m+1} \\ \mu_{m+2} \\ \vdots \\ \mu_{m+p} \end{pmatrix} \quad \text{con } \mu_i = 0 \text{ si } i \notin B(x_*) \\ -\lambda_* &= \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{pmatrix} \quad -\lambda_i = t_i - t'_i, \quad i \in I, \end{aligned}$$

en (5.8) resulta

$$\begin{aligned} \nabla f(x_*) &= - \sum_{i \in I} (\lambda_*)_i \nabla c_i(x_*) + \sum_{i \in D} (\mu_*)_i \nabla c_i(x_*) \\ \nabla f(x_*) + \sum_{i \in I} (\lambda_*)_i \nabla c_i(x_*) - \sum_{i \in D} (\mu_*)_i \nabla c_i(x_*) &= 0 \\ \nabla_x \ell(x_*, \lambda_*, \mu_*) &= 0, \end{aligned}$$

que es la condición (a).

Obviamente  $c_i(x_*) \geq 0, i \in D$  y  $c_i(x_*) = 0, i \in I$  pues  $x_*$  es factible para el problema PNL, entonces se tiene la condición (b).

Ahora bien si  $i \in B(x_*)$ , entonces  $c_i(x_*) = 0$  y por lo tanto se tiene la condición (c),  $(\mu_*)_i c_i(x_*) = 0$ .

Si  $i \notin B(x_*)$ , entonces  $c_i(x_*) > 0$ , y de acuerdo a la definición del vector  $\mu_*$ , esas componentes son ceros. Por los tanto  $(\mu_*)_i c_i(x_*) = 0$ , y se concluye que la condición de complementaridad se cumple para todo  $i \in D$ .

Finalmente, los multiplicadores  $(\mu_*)_i$  se han definido no negativos, y por lo tanto se cumple la condición (d). ■

## 5.3 Condiciones de optimalidad

Para aplicar el lema 5.2 hay que determinar si  $Z_2(x_*) = \emptyset$ . Si todas las funciones son de clase  $C^1$ ,  $Z_2(x_*) = \emptyset$  es una condición necesaria y suficiente para la existencia de los multiplicadores.

Introducimos ahora el concepto de condiciones calificadoras de primer y segundo orden. estas condiciones son propias de la región de factibilidad y pueden ser requeridas para que un punto sea solución del problema.

**Notación:** Se dice que  $PNL$  es  $C^k$  significando que  $f$  y  $c_i$ ,  $i \in I \cup D$  son clase  $C^k(H)$ .

### 5.3.1 Condición necesaria de primer orden

#### Definición 5.3 (Condición calificadora de primer orden)

Dado el problema  $PNL$ , se dice que una condición calificadora de primer orden de las restricciones se cumple en un punto factible  $x_*$  si para todo  $z$  tal que

$$\langle \nabla c_i(x_*), z \rangle = 0, \quad i \in I, \quad (5.9)$$

$$\langle \nabla c_j(x_*), z \rangle = 0, \quad j \in B(x_*), \quad (5.10)$$

entonces existe un arco factible de clase  $C^1$ ,  $\mathcal{A} : [0, \tau) \rightarrow H$  partiendo de  $x_*$  y tangente a  $z$ . Esto es

$$\mathcal{A}(0) = x_*, \quad \mathcal{A}'(0) = z.$$

Observemos que la función objetivo no interviene en la definición, dado que la condición calificadora se refiere sólo a las restricciones.

**Teorema 5.1 (Condición necesaria de primer orden)** *Asumiendo que  $PNL$  es  $C^1$ ,  $x_* \in H$  tal que la condición calificadora de primer orden se satisface en  $x_*$ . Entonces una condición necesaria para que  $x_*$  sea solución del  $PNL$  es que se cumplan las condiciones  $KKT$ .*

**Demostración:** Ver [13], [22]. ■

La condición calificadora es una propiedad hermosa, pero no muy útil porque es difícil verificar en la práctica. Por ejemplo  $\mathcal{A}(t)$  puede ser la solución de un problema de ecuaciones diferenciales.

### 5.3.2 Condición necesaria de segundo orden

Para distinguir los máximos de mínimos hay que tener en cuenta la curvatura de  $f$  y de las restricciones.

**Definición 5.4 ( Condición calificadoradora de segundo orden)**

Sea  $x_*$  un punto factible de PNL y sea  $z \in H$  tal que

$$\begin{aligned} \langle z, \nabla c_i(x_*) \rangle &= 0 & i \in I \\ \langle z, \nabla c_j(x_*) \rangle &= 0 & j \in B(x_*). \end{aligned}$$

Se dice que la condición calificadoradora de segundo orden de las restricciones se mantiene en  $x_*$  si existe un arco  $\mathcal{A}$  factible, de clase  $C^1$  en un entorno de 0, partiendo de  $x_*$  tal que

(a)  $\mathcal{A}(t)$  es tangente a  $z$ , esto es  $\mathcal{A}(0) = x_*$  y  $\mathcal{A}'(0) = z$ .

**Punto crucial de la definición**

(b)  $c_j(\mathcal{A}(t)) = 0, \quad t \in [0, \tau), \quad j \in B(x_*).$

**Teorema 5.2 (Condición necesaria de segundo orden)** Sea  $x_*$  una solución de PNL  $\in C^2$ . Asumiendo que la condición calificadoradora de segundo orden se verifica en  $x_*$ , entonces para todo  $\lambda_*, \mu_*$  tal que  $\nabla_x \ell(x_*, \lambda_*, \mu_*) = 0$  con  $(\mu_*)_i = 0, i \notin B(x_*)$  se tiene necesariamente que

$$\langle z, \nabla_x^2 \ell(x_*, \lambda_*, \mu_*) z \rangle_H \geq 0,$$

para todo  $z$  que satisface las condiciones (a) y (b) de la definición 5.4.

**Demostración:** Ver [13, 22]. ■

### 5.3.3 Condiciones suficientes

**Teorema 5.3 (Condición suficiente)**

Supongamos que PNL es clase  $C^2$ . Las condiciones suficientes para que  $x_*$  sea un mínimo local estricto de PNL son

i) Se verifiquen las condiciones KKT.

ii) Para todo sucesión  $\{x_k\}$  tal que  $x_k$  es factible y  $x_k \neq x_*$ ,  $x_k \rightarrow x_*$  y

$$\begin{aligned} \lim_{k \rightarrow \infty} \left\langle \nabla f(x_k), \frac{x_k - x_*}{\|x_k - x_*\|} \right\rangle_H &= 0 \Rightarrow \\ \limsup_{k \rightarrow \infty} \left\langle \frac{(x_k - x_*)}{\|x_k - x_*\|}, \nabla^2 \ell(x_*, \lambda_*, \mu_*) \frac{(x_k - x_*)}{\|x_k - x_*\|} \right\rangle_H &> 0. \end{aligned}$$

**Demostración:** Ver [13, 22]. ■

### 5.3.4 Aplicación

**Teorema 5.4 (Caso de dimensión finita)** *Sea PNL con  $H = \mathbb{R}^n$  y supongamos que PNL es  $C^2$ . Las siguientes condiciones son suficientes para que  $x_*$  sea una solución de PNL.*

i) *Se satisfacen las condiciones KKT.*

ii)  *$z^T \nabla^2 \ell(x_*, \lambda_*, \mu_*) z > 0$  siempre que*

$$\nabla c_i(x_*)^T z = 0, \quad \forall i \in I. \quad (5.11)$$

$$\nabla c_j(x_*)^T z = 0, \quad j \in M(x_*) = \{j : (\mu_*)_j > 0\}. \quad (5.12)$$

$$\nabla c_j(x_*)^T z \geq 0, \quad j \in B(x_*) - M(x_*). \quad (5.13)$$

**Demostración:** Supongamos que  $x_*$  no es un mínimo local estricto del problema PNL, entonces existe una sucesión  $\{x_k\}$  tal que

$x_k$  es factible.

$x_k \neq x_*$  y  $x_k \rightarrow x_*$ .

$f(x_k) \leq f(x_*)$ .

Definimos la subsucesión

$$\eta_{k_j} = \frac{x_{k_j} - x_*}{\|x_{k_j} - x_*\|},$$

que satisface  $\eta_{k_j} \rightarrow \bar{\eta} \neq 0$ .

El punto de acumulación  $\bar{\eta}$  verifica las siguientes propiedades.

$$\nabla c_j(x_*)^T \bar{\eta} \geq 0, \quad \forall j \in B(x_*). \quad (5.14)$$

$$\nabla c_i(x_*)^T \bar{\eta} = 0, \quad \forall i \in I. \quad (5.15)$$

$$\nabla f(x_*)^T \bar{\eta} \leq 0. \quad (5.16)$$

$$\nabla c_j(x_*)^T \bar{\eta} = 0, \quad \forall j \in M(x_*) = \{j : \mu_j > 0\} \subset B(x_*). \quad (5.17)$$

La primera propiedad es cierta pues para todo  $j \in B(x_*)$ ,

$$\frac{c_j(x_* + \|x_{k_j} - x_*\| \eta_{k_j}) - c_j(x_*)}{\|x_{k_j} - x_*\|} \geq 0.$$

Haciendo tender  $k_j \rightarrow \infty$  se tiene (5.14).

De la misma manera (5.15) se cumple pues para  $i \in I$ ,

$$\frac{c_i(x_\star + \|x_{k_j} - x_\star\|\eta_{k_j}) - c_i(x_\star)}{\|x_{k_j} - x_\star\|} = 0,$$

y haciendo tender  $k_j \rightarrow \infty$  se tiene que la igualdad.

Dado que  $f(x_{k_j}) \leq f(x_\star)$ ,

$$\frac{f(x_\star + \|x_{k_j} - x_\star\|\eta_{k_j}) - f(x_\star)}{\|x_{k_j} - x_\star\|} \leq 0.$$

Tomando límite para  $k_j \rightarrow \infty$ , se tiene  $\nabla f(x_\star)^T \bar{\eta} \leq 0$ , que es (5.16).

Finalmente veamos que se cumple (5.17).

$$\nabla f(x_\star) = \nabla \ell(x_\star, \lambda_\star, \mu_\star) - \sum_{i \in I} (\lambda_\star)_i \nabla c_i(x_\star) + \sum_{j \in D} (\mu_\star)_j \nabla c_j(x_\star).$$

Como se cumplen las condiciones KKT, el primer sumando es cero. Multiplicando escalarmente por  $\bar{\eta}$  y usando (5.15) y (5.16) resulta que

$$\begin{aligned} 0 \geq \nabla f(x_\star)^T \bar{\eta} &= - \sum_{i \in I} (\lambda_\star)_i \nabla c_i(x_\star)^T \bar{\eta} + \sum_{j \in D} (\mu_\star)_j \nabla c_j(x_\star)^T \bar{\eta} \\ &= \sum_{j \in D} (\mu_\star)_j \nabla c_j(x_\star)^T \bar{\eta}. \end{aligned}$$

Entonces

$$\sum_{j \notin M(x_\star)} (\mu_\star)_j \nabla c_j(x_\star)^T \bar{\eta} + \sum_{j \in M(x_\star)} (\mu_\star)_j \nabla c_j(x_\star)^T \bar{\eta} \leq 0. \quad (5.18)$$

Obsevamos que para que la desigualdad (5.18) se mantenga debe ser

$$\sum_{j \in M(x_\star)} (\mu_\star)_j \nabla c_j(x_\star)^T \bar{\eta} \leq 0.$$

Pero, para  $j \in M(x_\star)$ ,  $(\mu_\star)_j > 0$  y de acuerdo a (5.15)  $\nabla c_j(x_\star)^T \bar{\eta} \leq 0$ . Por lo tanto para todo  $j \in M(x_\star)$  debe ser  $\nabla c_j(x_\star)^T \bar{\eta} = 0$ .

En consecuencia el punto de acumulación  $\bar{\eta}$  satisface las propiedades (5.11), (5.12), (5.13) de la hipótesis ii). Sin embargo la matriz Hessiana  $\nabla^2 \ell(x_\star, \lambda_\star, \mu_\star)$  no es definida positiva. En efecto

$$\ell(x_{k_j}) - \ell(x_\star) = f(x_{k_j}) - f(x_\star) \leq 0$$

$$\nabla \ell(x_\star, \lambda_\star, \mu_\star)^T (x_{k_j} - x_\star) + \frac{1}{2} (x_{k_j} - x_\star)^T \nabla^2 \ell(x_\star, \lambda_\star, \mu_\star) (x_{k_j} - x_\star) + o(\|x_{k_j} - x_\star\|^2) \leq 0.$$

Por la hipótesis i), el primer sumando se anula. Dividiendo por  $\|x_{k_j} - x_\star\|^2$

$$(x_{k_j} - x_\star)^T \nabla^2 \ell(x_\star, \lambda_\star, \mu_\star) (x_{k_j} - x_\star) + o(\|x_{k_j} - x_\star\|^2) \leq 0,$$

tomando límite para  $k_j \rightarrow \infty$  se tiene

$$\bar{\eta}^T \nabla^2 \ell(x_\star, \lambda_\star, \mu_\star) \bar{\eta} \leq 0,$$

que es una contradicción a la hipótesis i). Por lo tanto  $x_\star$  tiene que ser un mínimo local de PNL. ■

### 5.3.5 Regularidad

Regularidad es una propiedad muy importante a causa de la relación que existe entre ella y las condiciones calificadoras de primer y segundo orden. Ya se dijo que en práctica éstas últimas son muy difícil de verificar, sin embargo, verificar regularidad es casi elemental.

#### Definición 5.5 (Punto regular)

Asumiendo que PNL es clase  $C^1$ , sea  $x_\star$  in punto factible de PNL. Se dice que  $x_\star$  es regular si el conjunto

$$\{\nabla c_i(x_\star), i \in I, \nabla c_j(x_\star), j \in B(x_\star)\}$$

es linealmente independiente.

**Teorema 5.5 (Regularidad)** Sea  $x_\star$  un punto regular de PNL.

(a) Si PNL es de clase  $C^1$ , entonces  $x_\star$  satisface la condición calificadora de primer orden.

(a) Si PNL es de clase  $C^2$ , entonces  $x_\star$  satisface la condición calificadora de segundo orden.

**Demostración:** Ver [13], [22]. ■



## Clase 6

# Algunos métodos numéricos para PNL

Entre los métodos numéricos para resolver problemas de PNL, mencionamos:

1. *Métodos de penalización*
2. *Método de los multiplicadores*
3. *Métodos de programación cuadrática sucesiva.*

Los métodos de penalización así como el método de los multiplicadores se basa en la construcción de una sucesión de subproblemas muy sencillos, tales como problemas sin restricciones, con restricciones de caja, etc.

Cuando la solución del problema, se estima resolviendo una sucesión de subproblemas sin restricciones, la técnica se conoce con el nombre de *SUMT*, *sequential unconstrained minimization techniques*: técnicas secuenciales de minimización sin restricciones [13]. Sin embargo, el objetivo no sólo es usar técnicas ya conocidas, sino que mediante una penalización, las restricciones que no se satisfacen son sancionadas de modo que las funciones que definen estas restricciones son eliminadas, aumentando la función objetivo introduciendo un término de penalización que involucra tales restricciones. Se distinguen dos tipos de penalización:

1. **Penalización externa** o estrategia de punto exterior. Esta técnica consiste en aumentar la función objetivo en un término cuyo costo aumenta con la violación o error que se comete en las restricciones. En general la solución del problema penalizado está fuera del conjunto factible, pero se aproxima a él cuando el término de penalización es muy grande.

2. **Penalización interna** o estrategia de punto interior. Los métodos de penalización interna o métodos de barrera son propuestos para problemas de optimización con restricciones de desigualdad.

El método de programación cuadrática sucesiva es uno de los más populares para resolver el problema general de PNL, quizá debido a que tiene un comportamiento similar al método de Newton para el caso sin restricciones.

Finalmente no debemos dejar de mencionar los métodos de gradiente proyectado [31], [32] y de gradiente reducido [33], [34], [19], [1].

## 6.1 Método de Newton

Consideremos el problema con restricciones de igualdad

$$(MRI) \equiv \begin{cases} \min & f(x) \\ \text{s.a} & C(x) = 0, \end{cases}$$

donde  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $C : \mathbb{R}^n \rightarrow \mathbb{R}^m$ , con  $C(x) = (c_1(x), \dots, c_m(x))^T$  y  $c_i : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $i \in I = \{1, \dots, m\}$ .

Dada una solución  $x_*$  de MRI, bajo ciertas condiciones la teoría de KKT dice que existen  $\lambda_* \in \mathbb{R}^m$  tal que

$$\nabla \ell(x_*, \lambda_*) = 0 \tag{6.1}$$

$$C(x_*) = 0. \tag{6.2}$$

donde

$$\ell(x, \lambda) = f(x) + \lambda^T C(x) = f(x) + \sum_{i=1}^m \lambda_i c_i(x)$$

$$\nabla \ell(x, \lambda) = \nabla f(x) + \nabla C(x) \lambda \in \mathbb{R}^n$$

$$\nabla C(x) = [\nabla c_1(x) | \nabla c_2(x) | \dots | \nabla c_m(x)] \in \mathbb{R}^{n \times m}.$$

Las condiciones necesarias de primer orden definen un sistema no lineal de  $n + m$  ecuaciones con  $n + m$  variables:  $x$  y  $\lambda$ , llamado **sistema extendido**,

$$\nabla \ell(x, \lambda) = \begin{pmatrix} \nabla_x \ell(x, \lambda) \\ C(x) \end{pmatrix} = 0.$$

El objetivo es establecer condiciones para poder aplicar el método de Newton al sistema extendido y de ese modo obtener un punto estacionario.

**Hipótesis estandar para el método de Newton**

1.  $x_*$  es una solución local de MRI y  $\lambda_*$  su multiplicador asociado tal que  $\nabla\ell(x_*, \lambda_*) = 0$ .
2.  $\nabla^2 f$ ,  $\nabla^2 c_i$  son Lipschitz continuas en un entorno de  $x_*$ .
3.  $\nabla^2\ell(x_*, \lambda_*)$  es no singular.

Observemos que

$$\nabla^2\ell(x_*, \lambda_*) = \begin{pmatrix} \nabla_x^2\ell(x_*, \lambda_*) & \nabla C(x_*) \\ \nabla C(x_*)^T & 0 \end{pmatrix},$$

simétrica y *nunca definida positiva*. El bloque abajo y a la derecha garantiza que nunca se obtiene un máximo o un mínimo de  $\ell$  en  $x$  y  $\lambda$ .

**Algoritmo 6.1 (Método de Newton para MRI)**

Sea  $x_c$  el iterado actual y  $\lambda_c$  el multiplicador asociado.

Repetir "hasta convergencia"

1. Resolver para  $s_c$  y  $\Delta\lambda_c$  el sistema lineal

$$\begin{pmatrix} \nabla_x^2\ell(x_c, \lambda_c) & \nabla C(x_c) \\ \nabla C(x_c)^T & 0 \end{pmatrix} \begin{pmatrix} s \\ \Delta\lambda \end{pmatrix} = - \begin{pmatrix} \nabla_x\ell(x_c, \lambda_c) \\ C(x_c) \end{pmatrix}.$$

2. Hacer

$$\begin{pmatrix} x_+ \\ \lambda_+ \end{pmatrix} = \begin{pmatrix} x_c \\ \lambda_c \end{pmatrix} + \begin{pmatrix} s_c \\ \Delta\lambda_c \end{pmatrix}.$$

La expresión "hasta convergencia" significa que  $x_c, \lambda_c$  es un punto de KKT o bien una muy buena aproximación de él. Esto significa que en cada iteración hay que verificar si

$$\|\nabla_x\ell(x_c, \lambda_c)\| \leq \tau_1 \quad y \quad \|C(x_c)\| \leq \tau_2.$$

Esto es equivalente a verificar

$$\|\nabla_x\ell(x_c, \lambda_c)\| + \|C(x_c)\| \leq \tau_3,$$

donde  $\tau_1$ ,  $\tau_2$  y  $\tau_3$  son tolerancias prefijadas.

**Definición 6.1 (Punto no singular)**

Cualquier punto  $x_* \in \mathbb{R}^n$  que satisface  $\nabla\ell(x_*, \lambda_*) = 0$  se dice no singular para el problema MRI si  $\nabla^2\ell(x_*, \lambda_*)$  es no singular.

**Teorema 6.1** Sean  $x_*$ ,  $\lambda_*$  tal que  $\nabla\ell(x_*, \lambda_*) = 0$

i)  $x_*$  es un punto no singular

ii)  $x_*$  es un punto regular

iii)  $\lambda_*$  es único.

Entonces i)  $\Rightarrow$  ii)  $\Rightarrow$  iii).

**Demostración:**

i)  $\Rightarrow$  ii) Supongamos que  $x_*$  no singular. Entonces la matriz  $\nabla^2(\ell(x_*, \lambda_*))$  es no singular, entonces las últimas  $m$  columnas son linealmente independientes y por lo tanto  $x_*$  es regular.

ii)  $\Rightarrow$  iii) Asumimos que  $x_*$  es regular. Entonces  $\nabla C(x_*) \in \mathbb{R}^{n \times m}$  tiene  $m$  columnas linealmente independientes. Consideramos

$$\begin{aligned}\nabla\ell(x_*, \lambda_*) &= \nabla f(x_*) + \nabla C(x_*)\lambda_* = 0 \\ \nabla C(x_*)\lambda_* &= -\nabla f(x_*) \\ \nabla C(x_*)^T \nabla C(x_*)\lambda_* &= -\nabla(x_*)^T \nabla f(x_*)\end{aligned}\quad (6.3)$$

que son las ecuaciones normales. Como  $\nabla C(x_*) \in \mathbb{R}^{n \times m}$  tiene rango completo, la matriz  $\nabla C(x_*)^T \nabla C(x_*)$  es no singular. Por lo tanto el sistema (6.3) tiene solución única,

$$\lambda_* = \lambda_{CM} = -(\nabla C(x_*)^T \nabla C(x_*))^{-1} \nabla(x_*)^T \nabla f(x_*).$$

■

La importancia de la no singularidad de un punto yace en su relación con la condición de suficiencia. *Un punto estacionario no singular satisface las condiciones suficiente de segundo orden* como muestra el siguiente teorema.

**Teorema 6.2** Sea  $x_*$  una solución local de MRI. Asumiendo que  $x_*$  es un punto regular, entonces  $x_*$  es no singular si y sólo si

$$\eta^T \nabla_x^2 \ell(x_*, \lambda_*) \eta > 0, \quad \forall \eta \neq 0. \quad (6.4)$$

$$\nabla C(x_*)^T \eta = 0 \quad (6.5)$$

La demostración del teorema es una aplicación directa del siguiente lema.

**Lema 6.1** Sean  $A \in \mathbb{R}^{n \times n}$  simétrica,  $B \in \mathbb{R}^{n \times m}$  una matriz de rango  $m$ . Si  $v^T Av \geq 0$  para todo  $v \in \mathbb{R}^n$  tal que  $B^T v = 0$ , entonces

$$v^T Av > 0, \quad v \neq 0 \iff \begin{pmatrix} A & B \\ B^T & O \end{pmatrix}$$

es no singular.

**Demostración:**  $\Leftarrow$ ) Supongamos que

$$\begin{pmatrix} A & B \\ B^T & O \end{pmatrix}$$

es no singular y que existe  $v_0 \neq 0$  tal que  $v_0^T Av_0 = 0$  con  $B^T v_0 = 0$ . Obviamente  $v_0$  es una solución del problema

$$\min_{B^T v = 0} \frac{1}{2} v^T Av$$

Aplicamos la teoría de KKT a este problema. Sea

$$\ell(v, \xi) = \frac{1}{2} v^T Av + \xi^T (B^T v).$$

De las condiciones de KKT

$$\begin{aligned} \nabla_v \ell(v_0, \xi_0) &= 0 \\ B^T v_0 &= 0 \end{aligned}$$

resulta

$$\begin{pmatrix} A & B \\ B^T & O \end{pmatrix} \begin{pmatrix} v_0 \\ \xi_0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Luego  $v_0 = 0$  que contradice la hipótesis.

$\Rightarrow$ ) Suponemos que

$$\begin{pmatrix} A & B \\ B^T & O \end{pmatrix}$$

es singular. Entonces existen  $v_0 \in \mathbb{R}^n$ ,  $\xi_0 \in \mathbb{R}^m$ ,  $v_0 \neq 0$  o  $\xi_0 \neq 0$  tal que

$$\begin{pmatrix} A & B \\ B^T & O \end{pmatrix} \begin{pmatrix} v_0 \\ \xi_0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Esto es

$$\begin{aligned} Av_0 + B\xi_0 &= 0 \\ B^T v_0 &= 0. \end{aligned}$$

Multiplicando la primer ecuación por  $v_0^T (\neq 0)$ ,

$$v_0^T Av_0 + v_0^T B\xi_0 = 0$$

como  $B^T v_0 = 0$  resulta  $v_0^T Av_0 = 0$  con  $v_0 \neq 0$ , que contradice la hipótesis. ■

**Demostración del teorema.** La demostración del teorema 6.2 sigue del hecho que

$$\nabla^2 \ell(x_*, \lambda_*) = \begin{pmatrix} \nabla_x^2 \ell(x_*, \lambda_*) & \nabla C(x_*) \\ \nabla C(x_*)^T & O \end{pmatrix}$$

y el lema vale con  $A = \nabla_x^2 \ell(x_*, \lambda_*)$  y  $B = \nabla C(x_*)$ . ■

### 6.1.1 Ejercicios

**Ejercicio 6.1** Sea  $x_*$  una solución local del problema MRI. Demuestre que:

- a) Si  $x_*$  es un punto no singular y satisface las condiciones necesarias de primer y segundo orden entonces  $x_*$  también satisface la condición suficiente de segundo orden.
- b) Si  $x_*$  es regular y satisface condiciones suficientes de segundo orden entonces  $x_*$  es un punto no singular.

**Ejercicio 6.2** Sea  $x_*$  un punto regular, mínimo local del problema MRI. Pruebe que  $\nabla f(x_*)$  es ortogonal al espacio tangente  $\mathcal{N}(\nabla C(x_*)^T)$ .

## 6.2 Programación cuadrática sucesiva

Uno de los métodos más populares para resolver problemas de programación no lineal

$$(PNL) \equiv \begin{cases} \text{Minimice} & f(x) \\ \text{s.a} & c_i(x) = 0 \quad i \in I \\ & c_i(x) \geq 0 \quad i \in D. \end{cases}$$

es el método de **programación cuadrática sucesiva (PCS)**, debido a Wilson, 1963.

En los métodos tipo SUMT, como los métodos de penalización y el métodos de los multiplicadores, el problema con restricciones se transforma en un problema sin restricciones. La estrategia de PCS es diferente.

El método se puede describir como sigue: sea  $x_*$  un punto estacionario de PNL. Si

$x_c$  es la aproximación actual de  $x_*$ , el método halla un paso  $s_c (= s_c^{PC}) \in \mathbb{R}^n$  resolviendo aproximadamente el siguiente problema cuadrático

$$(PC) \equiv \begin{cases} \min & q_c(s) \\ \text{s.a} & \nabla c_i(x_c)^T s + c_i(x_c) = 0 \quad i \in I \\ & \nabla c_j(x_c)^T s + c_j(x_c) \geq 0 \quad j \in D, \end{cases}$$

siendo  $q_c(s)$  un modelo cuadrático de la función de Lagrange asociada con PNL

$$\ell(x, \lambda, \mu) = f(x) + \sum_{i \in I} \lambda_i c_i(x) - \sum_{j \in D} \mu_j c_j(x),$$

y la minimización se efectúa sobre una aproximación lineal de las restricciones.

Sin pérdida de generalidad asumimos  $D = \emptyset$ . Esto significa que el modelo es

$$(PC) \equiv \begin{cases} \min & q_c(s) \\ \text{s.a} & \nabla c_i(x_c)^T s + c_i(x_c) = 0 \quad i \in I. \end{cases}$$

### 6.2.1 Construcción del modelo cuadrático

En la literatura encontramos que hay dos formas de construir el modelo cuadrático de  $\ell(x_c + s, \lambda_c)$  alrededor del iterado actual  $x_c$ .

$$q_c(s) = \frac{1}{2} s^T H_c s + \nabla \ell_c^T s + \ell_c \quad (6.6)$$

$$q_c(s) = \frac{1}{2} s^T H_c s + \nabla f_c^T s + \ell_c \quad (6.7)$$

donde  $H_c = \nabla^2 \ell(x_c, \lambda_c)$  o bien una aproximación de la matriz Hessiana de  $\ell$  en el iterado actual.

Veamos que significa elegir un modelo u otro.

1. Una primera elección fue

$$q_c(s) = \frac{1}{2} s^T H_c s + \nabla f_c^T s + f_c \quad \text{con } H_c = \nabla^2 f_c.$$

Con esta elección no hay ninguna posibilidad de probar convergencia, ver Fletcher [14].

Analicemos las condiciones de primer orden para el subproblema PC. La función de Lagrange es

$$\varphi(s, \xi) = q_c(s) + \xi^T (\nabla C_c^T s + C_c)$$

y las condiciones de KKT,

$$\begin{aligned} H_c s + \nabla f_c + \nabla C_c \xi &= 0 \\ \nabla C_c^T s + C_c &= 0 \end{aligned}$$

o

$$H_c s + \nabla C_c \xi = -\nabla f_c \quad (6.8)$$

$$\nabla C_c^T s = -C_c \quad (6.9)$$

$$\begin{pmatrix} H_c & \nabla C_c \\ \nabla C_c^T & 0 \end{pmatrix} \begin{pmatrix} s \\ \xi \end{pmatrix} = - \begin{pmatrix} \nabla f_c \\ C_c \end{pmatrix}.$$

Veamos que significado tiene el multiplicador  $\xi$ . Dado que  $s$  es la corrección del iterado actual, para obtener una mejor (o no) aproximación de la solución  $x_*$ , es deseable y lógico que el lugar de  $\xi$  fuera ocupado por la corrección del multiplicador  $\lambda_c$ .

Si en la ecuación (6.8) de las condiciones de KKT restamos a ambos miembros el vector  $\nabla C_c \lambda_c$  resulta

$$H_c s + \nabla C_c \xi - \nabla C_c \lambda_c = -\nabla f_c - \nabla C_c \lambda_c$$

que junto con (6.9) permiten escribir

$$\begin{aligned} H_c s + \nabla C_c (\xi - \lambda_c) &= -\nabla \ell(x_c, \lambda_c) \\ \nabla C_c^T s &= -C_c. \end{aligned}$$

Haciendo  $\xi - \lambda_c = \Delta \lambda_c$ , esto es  $\xi$  debe ser  $\lambda_+$ , resulta el sistema

$$\begin{pmatrix} H_c & \nabla C_c \\ \nabla C_c^T & 0 \end{pmatrix} \begin{pmatrix} s \\ \Delta \lambda \end{pmatrix} = - \begin{pmatrix} \nabla \ell(x_c, \lambda_c) \\ C_c \end{pmatrix}.$$

Ahora  $s$  es la corrección del iterado  $x_c$  y  $\Delta \lambda$  es la corrección del multiplicador, y se puede aplicar el método de Newton si la matriz

$$\begin{pmatrix} H_c & \nabla C_c \\ \nabla C_c^T & 0 \end{pmatrix}$$

fuera la matriz Jacobiana del lado derecho. Por lo tanto se obtiene (6.6).



2. Ahora bien si  $H_c \approx \nabla^2 \ell_c$  y si  $\nabla \ell_c$  se reduce a  $\nabla f_c$  se tiene  $\lambda_+$ , que es la actualización del multiplicador, elección que corresponde al modelo (6.7). De hecho estamos resolviendo

$$\begin{pmatrix} H_c & \nabla C_c \\ \nabla C_c^T & 0 \end{pmatrix} \begin{pmatrix} s \\ \lambda - \lambda_c \end{pmatrix} = - \begin{pmatrix} \nabla f_c + \nabla C_c \lambda_c \\ C_c \end{pmatrix}.$$

es decir se resuelve

$$\begin{pmatrix} H_c & \nabla C_c \\ \nabla C_c^T & 0 \end{pmatrix} \begin{pmatrix} s \\ \lambda \end{pmatrix} = - \begin{pmatrix} \nabla f_c \\ C_c \end{pmatrix}.$$

**Conclusión:**

1. Los dos modelos (6.6) y (6.7) dan el mismo paso  $s$ .
2. Los multiplicadores no son los mismos. El modelo (6.7) dá el verdadero multiplicador,  $\lambda_+$ , mientras que el modelo (6.6) proporciona la corrección  $\Delta \lambda_c$ .

**6.2.2 El algoritmo de programación cuadrática sucesiva**

El siguiente es un esquema de una iteración del algoritmo basado en PCS.

**Algoritmo 6.2 (Iteración de PCS)**  
 Dado  $x_c \in \mathbb{R}^n$

**Paso 1.** Hallar un paso  $s_c$  como solución aproximada de

$$(PC) \equiv \begin{cases} \min & q_c(s) \\ \text{s.a} & \nabla c_i(x_c)^T s + c_i(x_c) = 0 \quad i \in I \end{cases}$$

**Paso 2.** Hacer  $x_+ = x_c + s_c$

**Paso 3:** Actualizar o corregir el multiplicador.

**Ventajas del método**

Son varias las razones que hacen a PCS una estrategia popular:

1. Se comporta de la misma manera que el método de Newton para el caso sin restricciones: bajo las hipótesis estándar, si  $H_c = \nabla^2 \ell_c$ , converge local y  $q$ -cuadráticamente y si  $H_c$  es actualizado por un método secante la convergencia es  $q$ -superlineal.
2. Es sencillo extender el algoritmo incluyendo restricciones de desigualdad vía algún método de restricciones activas.
3. Existen códigos excelentes para resolver el subproblema cuadrático.

### Desventajas del método

Las desventajas del método son las mismas que las del método de Newton para el caso sin restricciones:

1. No converge globalmente sin una estrategia de globalización.
2. El subproblema puede no tener solución si  $H_c$  no es definida positiva en  $\mathcal{N}(\nabla C_c^T)$  o si  $\nabla C_c^T$  no tiene rango completo.  
La *existencia* de solución está garantizada si  $\text{rang}(\nabla C_c^T) = m$ , mientras que la *unicidad* si  $H_c$  es definida positiva en  $\mathcal{N}(\nabla C_c^T)$ .

### 6.2.3 Solución del problema cuadrático

En cada iteración del algoritmo, se debe resolver un subproblema de programación cuadrática

$$(PC) \equiv \begin{cases} \min & q_c(s) \\ \text{s.a} & \nabla c_i(x_c)^T s + c_i(x_c) = 0 \quad i \in I. \end{cases}$$

Veamos como hallar  $s_c^{PC}$  y su multiplicador asociado  $\Delta \lambda_c^{PC}$ , cuando la solución existe y que hacer cuando la solución no existe.

Existe una variedad de algoritmos para hallar una solución aproximada de PC, algunos más eficientes que otros. Aquí vamos a presentar las formulaciones propuestas por

- *Dennis-Williamson* [37]: una implementación robusta que contempla todas las posibles situaciones.
- *Dennis-Maciel* [9, 21]: un algoritmo que usa direcciones conjugadas sobre  $\mathcal{N}(\nabla C_c^T)$ . Se adapta a problemas de gran porte.

Una dificultad que se puede encontrar al tratar de resolver PC es que la solución podría no existir, (si bien el PNL tiene solución). Entonces,

- 1) Si  $\text{rang}(\nabla C_c^T) < m$ , el algoritmo debe ser capaz de manejar situaciones en las cuales las restricciones son degeneradas.
- 2) Podría suceder que no haya ningún punto que sea *linealmente factible*, esto es, no existe  $s \in \mathbb{R}^n$  tal que  $\nabla C_c^T s + C_c = 0$ . Por lo tanto el subproblema no tiene solución !.

Para evitar este obstáculo, sea  $s_{CM}$  una solución del problema

$$\min_{s \in \mathbb{R}^n} \|\nabla C_c^T s + C_c\|_2,$$

y sea

$$\Theta_{MIN} = \nabla C_c^T s_{CM} + C_c \quad (6.10)$$

entonces, en lugar de considerar la restricción lineal  $\nabla C_c^T s + C_c = 0$ , consideramos

$$\begin{aligned} \nabla C_c^T s + C_c &= \Theta_{MIN} \\ \nabla C_c^T (s - s_{CM}) &= 0. \end{aligned} \quad (6.11)$$

De esta forma estamos pidiendo que el paso sea tan linealmente factible como sea posible. Entonces cuando el problema PC no tiene solución pues  $\nabla C_c^T$  no es de rango completo se resuelve el siguiente problema que llamamos **problema generalizado de programación cuadrática**,

$$(PCG) \equiv \begin{cases} \min & q_c(s) \\ \text{s.a} & \nabla C_c^T (s - s_{CM}) = 0, \end{cases}$$

y de esta forma garantizamos que el problema tiene solución.

**Conclusión:** Si PC no tiene solución porque no existe un paso  $s$  tal que  $\nabla C_c^T s + C_c = 0$ , esta restricción se relaja para obtener un subproblema que tenga solución.

- 3) Supongamos que la condición suficiente de segundo orden no se satisface, hecho que puede ocurrir si el iterado actual está lejos de la solución  $x_*$ . Si

$$z^T \nabla^2 \ell_c z > 0 \quad \text{con} \quad \nabla C_c^T z = 0$$

no se cumple, entonces existe una dirección  $d_{PC} \in \mathcal{N}(\nabla C_c^T)$  de curvatura negativa o nula. En este caso el problema no tiene solución.

En efecto, sea  $s = s_{CM} + \alpha d_{PC}$  con  $\alpha \in \mathbb{R}$ . Calculemos

$$\begin{aligned}\nabla C_c^T s + C_c &= \nabla C_c^T (s_{CM} + \alpha d_{PC}) + C_c \\ &= \nabla C_c^T s_{CM} + \alpha \nabla C_c^T d_{PC} + C_c \\ &= \nabla C_c^T s_{CM} + C_c \\ &= \Theta_{MIN}.\end{aligned}$$

Por lo tanto  $s = s_{CM} + \alpha d_{PC}$  satisface las restricciones. Ahora calculamos,

$$\begin{aligned}q_c(s) &= \frac{1}{2}(s_{CM} + \alpha d_{PC})^T H_c(s_{CM} + \alpha d_{PC}) + \nabla \ell_c^T (s_{CM} + \alpha d_{PC}) \\ &= q_c(s_{CM}) + \alpha[\nabla \ell_c^T d_{PC} + s_{CM}^T H_c d_{PC}] + \frac{1}{2}\alpha^2 d_{PC}^T H_c d_{PC} \\ &= q_c(s_{CM}) + \alpha \nabla q_c(s_{CM})^T d_{PC} + \frac{1}{2}\alpha^2 d_{PC}^T H_c d_{PC}\end{aligned}\quad (6.12)$$

Analizamos

$$\alpha^2 d_{PC}^T H_c d_{PC}.$$

(3a) Si  $d_{PC}^T H_c d_{PC} < 0$ , se puede elegir el signo de  $\alpha$  de modo que

$$\alpha \nabla q_c(s_{CM})^T d_{PC} = \alpha[\nabla \ell(x_c, \lambda_c)^T d_{PC} + s_{CM}^T H_c d_{CN}] \leq 0,$$

entonces si  $|\alpha| \rightarrow \infty$ , la cuadrática  $q_c(s) \rightarrow -\infty$ , para todo  $s$  de la forma  $s = s_{CM} + \alpha d_{PC}$ .

(3b) Si  $d_{PC}^T H_c d_{PC} = 0$ , es decir  $d_{PC}$  es una dirección de curvatura nula en el espacio tangente de las restricciones. Sin pérdida de generalidad podemos suponer que no existen direcciones de curvatura negativa en  $\mathcal{N}(\nabla C_c^T)$ , ya que se vió que en ese caso el problema no tiene solución.

En (6.12) la cuadrática se transforma en

$$q_c(s) = q_c(s_{CM}) + \alpha \nabla q_c(s_{CM})^T d_{PC}\quad (6.13)$$

(b1) Si  $\nabla q_c(s_{CM})^T d_{PC} \neq 0$ , como en (3a) se puede elegir el signo de  $\alpha$  de modo que si  $|\alpha| \rightarrow \infty$ , la cuadrática  $q_c(s) \rightarrow -\infty$ , y por lo tanto el problema no tiene solución.

(b2) Si  $\nabla q_c(s_{CM})^T d_{PC} = 0$ , para todas las direcciones de curvatura nula en  $\mathcal{N}(\nabla C_c^T)$  se tiene  $q_c(s) = q_c(s_{CM})$ , esto es  $d_{PC}$  es una dirección a lo largo de la cual  $q_c(s)$  no cambia.

Si  $s_{CM} = 0$ ,  $s = \alpha d_{PC}$  es una dirección de descenso siempre que el signo de  $\alpha$  sea elegido de modo que  $\alpha \nabla \ell_c^T d_{PC} < 0$ .

En el caso que  $d$  es una dirección de curvatura nula y  $\nabla \ell_c^T d_{PC} = 0$ , esto es  $d_{PC}$  no es una dirección de descenso, sino que a lo largo de  $d_{PC}$ , la función  $q_c$  permanece sin cambios. Pero pueden existir direcciones desde  $s_{CM}$ , que sean de descenso para  $q_c$ . Para determinar una de tales direcciones restringimos  $q_c(s)$  al espacio nulo de las restricciones y resolvemos

$$\begin{array}{ll} \min & q_c(s) \\ \text{s.a} & \\ & \nabla C_c^T s = 0. \end{array}$$

Sea  $\hat{s}$  la solución de este problema. Combinando  $\hat{s}$  con el paso  $s_{CM}$  y la dirección a lo largo de la cual  $q_c(s)$  permanece sin cambios se tiene un número infinito de soluciones del problema PCG de la forma

$$s(\alpha) = s_{CM} + \alpha d_{PC} + \hat{s}.$$

**Sumario:**

**I.** Si  $\text{rang}(\nabla C_c^T) < m \Rightarrow$

- Hallar  $s_{CM}$  como solución de  $\min \|\nabla C_c^T s + C_c\|_2$
- Definir

$$(PCG) \equiv \begin{cases} \min & q_c(s) \\ \text{s.a} & \nabla C_c^T (s - s_{CM}) = 0, \end{cases}$$

**II.** Sea  $s = s_{CM} + \alpha d_{PC}$ ,  $\alpha \in \mathbb{R}$ ,  $d_{PC} \in \mathcal{N}(\nabla C_c^T)$

1. Si  $d_{PC}$  es una dirección de curvatura negativa, es posible elegir el signo de  $\alpha$  tal que  $|\alpha| \rightarrow \infty$ , la cuadrática  $q_c(s) \rightarrow -\infty$ ,  
 (**★ además si  $\alpha \nabla \ell^T d_{PC} < 0$  entonces  $d_{PC}$  es una dirección de descenso ★**).

2. Si  $d_{PC}^T H_c d_{PC} = 0$ ,

- $\nabla q_c(s_{CN})^T d_{PC} \neq 0 \Rightarrow q_c(s) \rightarrow -\infty$ ,  
 (**★ además si  $\alpha \nabla \ell^T d_{PC} < 0$  entonces  $d_{PC}$  es una dirección de descenso ★**).
- $\nabla q_c(s_{CN})^T d_{PC} = 0 \Rightarrow q_c(s) = q_c(s_{CM})$  a lo largo de  $d_{PC}$ .  
 Para determinar si existe alguna dirección de descenso se halla  $\hat{s}$  como solución de

$$\begin{aligned} & \min && q_c(s) \\ & \text{s.a} && \\ & && \nabla C_c^T s = 0. \end{aligned}$$

$$s = s(\alpha) = s_{CM} + \alpha d_{PC} + \hat{s}.$$

#### 6.2.4 Formulación del algoritmo de Dennis y Williamson

Si  $s_{PC}$  es la solución del problema PCG y  $\Delta \lambda_{PC}$  el multiplicador asociado,  $s_{PC}$  y  $\Delta \lambda_{PC}$  satisfacen

$$\begin{pmatrix} H & \nabla C_c \\ \nabla C_c^T & 0 \end{pmatrix} \begin{pmatrix} s_{PC} \\ \Delta \lambda_{PC} \end{pmatrix} = \begin{pmatrix} -\nabla \ell_c \\ \nabla C_c \end{pmatrix}.$$

- **Paso 1.** Computa la factorización  $QR$  con pivoteo por columnas de  $\nabla C_c$ :

$$\nabla C_c = QR\Pi^T,$$

donde

$Q \in \mathbb{R}^{n \times n}$  es ortogonal

$R \in \mathbb{R}^{n \times m}$  es triangular superior

$\Pi \in \mathbb{R}^{m \times m}$  es la matriz permutación que describe el pivoteo por columnas.

Si  $\text{rang}(\nabla C_c) = r \leq m$ , entonces en la partición de las columnas de  $Q$  como  $Q = [Q_1|Q_2]$ , la matriz  $Q_1 \in \mathbb{R}^{n \times r}$  tiene  $r$ -columnas que forman una base de  $\text{Col}(\nabla C_c)$  y  $Q_2 \in \mathbb{R}^{n \times (n-r)}$  es tal que sus  $(n-r)$ -columnas generan  $\mathcal{N}(\nabla C_c^T)$ .

$$R = \begin{pmatrix} R_1 & R_2 \\ 0 & 0 \end{pmatrix}$$

con

$R_1 \in \mathbb{R}^{r \times r}$  es triangular superior y no singular

$R_2 \in \mathbb{R}^{r \times (n-r)}$ .

Así, el paso  $s \in \mathbb{R}^n$  se puede representar como la suma de dos componentes

$$s = \underbrace{Q_1 w_1}_{\in \text{Col}(\nabla C)} + \underbrace{Q_2 w_2}_{\in \text{Col}(\nabla C^T)},$$

donde  $w_1 \in \mathbb{R}^r$ ,  $w_2 \in \mathbb{R}^{n-r}$ .

Dado que  $r \leq m$ , el sistema  $\nabla C_c^T s = -C_c$  se resuelve en el sentido de los cuadrados mínimos. El algebra lineal numérica nos permite asegurar que resolver el sistema  $\nabla C_c^T s = -C_c$ , es equivalente a resolver

$$R_1^T w_1 = -(\Pi^T C)_r = \begin{pmatrix} \tilde{c}_1 \\ \tilde{c}_2 \\ \vdots \\ \tilde{c}_r \end{pmatrix}.$$

**Conclusión:** la componente  $w_1$  se puede determinar resolviendo un simple sistema triangular inferior.

Una vez que se tiene  $w_1$ , el paso  $s_{CM}$  se calcula como  $s_{CM} = Q_1 w_1 \equiv s_{LF}$

$$\begin{aligned}
\Theta_{MIN} &= \nabla C_c^T s_{CM} + C_c = \nabla C_c^T Q_1 w_1 + C_c \\
&= \Pi \begin{pmatrix} R_1^T & 0 \\ R_2^T & 0 \end{pmatrix} \begin{pmatrix} Q_1^T \\ Q_2^T \end{pmatrix} Q_1 w_1 + C_c \\
&= \Pi \begin{pmatrix} R_1^T Q_1^T \\ R_2^T Q_2^T \end{pmatrix} Q_1 w_1 + C_c = \Pi \begin{pmatrix} R_1^T \\ R_2^T \end{pmatrix} w_1 + \begin{pmatrix} C_r \\ C_{m-r} \end{pmatrix} \\
&= \Pi \left[ \begin{pmatrix} R_1^T \\ R_2^T \end{pmatrix} w_1 + \begin{pmatrix} (\Pi^T C)_r \\ (\Pi^T C)_{m-r} \end{pmatrix} \right] \\
&= \Pi \begin{pmatrix} R_1^T w_1 + (\Pi^T C)_r \\ R_2^T w_1 + (\Pi^T C)_{m-r} \end{pmatrix} = \Pi \begin{pmatrix} 0 \\ R_2^T w_1 + (\Pi^T C)_{m-r} \end{pmatrix}.
\end{aligned}$$

• **Paso 2.** Determina si existe una dirección de curvatura negativa o nula para  $H_c$  en  $\mathcal{N}(\nabla C_c^T)$ .

La matriz Hessiana restringida al espacio nulo es  $Q_2^T H_c Q_2$ .

Sea  $\Lambda_1$  el menor autovalor de  $Q_2^T H_c Q_2$  y  $v_1$  el autovector asociado.

Si  $\Lambda_1 < 0$ , entonces  $v_1$  es una dirección de curvatura negativa para  $H_c$ .

Si  $\Lambda_1 = 0$ , entonces  $v_1$  es una dirección de curvatura nula para  $H_c$ .

En cualquiera de los dos casos sea  $d = Q_2 v_1$ , donde  $Q_2$  es una base de  $\mathcal{N}(\nabla C_c^T)$ .

Sea  $\Lambda_1 = 0$  y  $\nabla q_c^T d = 0$ , entonces  $d$  no es una dirección de descenso para  $q$ . En este momento se debe determinar si existe algún autovector correspondiente a  $\Lambda_1 = 0$  que dé una dirección de descenso para  $q_c(s)$ . Si tal autovector existe, digamos  $v_i$  y  $\nabla \ell_c^T Q v_i \neq 0$ , definimos  $d = Q_2 v_i$ .

En otro caso el problema tiene infinitas soluciones y por lo tanto resolvemos

$$\min q_c(s) \quad s \in \mathcal{N}(\nabla C_c^T).$$

Sea  $\hat{s} = Q_2 v_2^*$  la solución, donde  $v_2^*$  resuelve  $Q_2^T H_c Q_2 v_2 = -Q_2^T \nabla \ell_c$ .

**Conclusión:** Si  $d$  es una dirección de curvatura negativa o nula, el problema PCG no tiene solución o bien tiene infinitas soluciones de la forma

$$s = s_{CM} + \hat{s} + \alpha d = s_{CM} + Q_2 v_2^* + \alpha Q_2 v_i = s_{CM} + Q_2 (v_2^* + \alpha v_i).$$

En cualquiera de las dos situaciones se dá por finalizado el algoritmo.

• **Paso 3.** Si se ha podido determinar que no existe dirección de curvatura  $\leq 0$  en  $\mathcal{N}(\nabla C_c^T)$ , se computa  $s_{PC} \in \mathcal{N}(\nabla C_c^T)$ .

$$H_c s + \nabla C_c \Delta \lambda_c = -\nabla \ell_c \quad (6.14)$$

$$s = q_1 w_1 + Q_2 w_2 = s_{CM} + Q_2 w_2 \quad (6.15)$$



entonces reemplazando (6.15) en (6.14) resulta

$$\begin{aligned} H_c(s_{CM} + Q_2 w_2) + \Delta \lambda_c &= -\nabla \ell_c \\ H_c Q_2 w_2 + \nabla C_c \Delta \lambda_c &= -(\nabla \ell_c + H_c s_{CM}) \\ Q^2 H_c Q_2 w_2 &= -(\nabla \ell_c + H_c s_{CM}). \end{aligned}$$

Esto es, se resuelve un sistema lineal en la componente  $w_2$ .

La solución es  $s_{PC} = s_{CM} + Q_2 w_2$ .

Falta ver como se determina  $\Delta \lambda_c$ , el multiplicador asociado a  $s_{CM}$ .

$$\begin{aligned} \nabla C_c \Delta \lambda &= -(\nabla \ell_c + H_c s_{PC}) \\ (Q_1 | Q_2) \begin{pmatrix} R_1 & R_2 \\ 0 & 0 \end{pmatrix} \underbrace{\Pi^T \Delta \lambda}_{\tilde{\Delta} \lambda} &= -(\nabla \ell_c + H_c s_{PC}) \\ (Q_1 | Q_2) \begin{pmatrix} R_1 & R_2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} (\tilde{\Delta} \lambda)_r \\ (\tilde{\Delta} \lambda)_{m-r} \end{pmatrix} &= -(\nabla \ell_c + H_c s_{PC}) \end{aligned} \quad (6.16)$$

Dado que las últimas  $(m-r)$  componentes de  $\tilde{\Delta} \lambda = \Pi^T \Delta \lambda$  corresponden a las columnas de  $\nabla C_c^T$  que son linealmente dependientes, podemos hacer estas componentes iguales a 0. Por lo tanto en (6.16) queda

$$\begin{aligned} (Q_1 R_1 | Q_2 R_2) \begin{pmatrix} (\tilde{\Delta} \lambda)_r \\ 0 \end{pmatrix} &= -(\nabla \ell_c + H_c s_{PC}) \\ Q_1 R_1 (\tilde{\Delta} \lambda)_r &= -(\nabla \ell_c + H_c s_{PC}) \\ R_1 (\tilde{\Delta} \lambda)_r &= -Q_1^T (\nabla \ell_c + H_c s_{PC}). \end{aligned}$$

Esto es,  $(\tilde{\Delta} \lambda)_r$  se obtiene como solución un sistema triangular superior, sencillo de resolver. Así

$$\tilde{\Delta} \lambda_c = \Pi^T \Delta \lambda_c = \begin{pmatrix} (\Pi^T \Delta \lambda_c)_r \\ 0 \end{pmatrix}.$$

Luego

$$\Delta \lambda_{PC} = \Pi \tilde{\Delta} \lambda_c.$$

### 6.2.5 Formulación del algoritmo de Dennis y Maciel

Dennis y Maciel, [9] proponen un algoritmo que resuelve el problema cuadrático de la siguiente manera:

1. Se determina un punto inicial linealmente factible, o el *más linealmente factible* si  $\text{rang}(\nabla C_c^T) < m$ .
2. Se reduce la cuadrática  $q_c(s)$  al espacio  $\mathcal{N}(\nabla C_c^T)$  transformando todo el problema en uno de dimensión considerablemente menor,  $n - r$ , donde  $r$  es la dimensión del espacio columna de  $\nabla C_c^T$ .
3. Se utilizan direcciones conjugadas.

Sea el problema cuadrático,

$$(PC) \equiv \begin{cases} \min & q_c(s) = \frac{1}{2}s^T H_c s + \nabla \ell_c^T s \\ \text{s.a} & \\ & \nabla C_c^T s + C_c = 0. \end{cases}$$

Consideremos primero el problema de factibilidad.

Supongamos  $C(x_c) = 0$  y  $\text{rang}(\nabla C_c^T) = r \leq m$ . Con la idea del método del gradiente reducido en mente consideremos una partición de  $\nabla C_c^T$ ,

$$\nabla C_c^T = [B|N]$$

donde  $B \in \mathbb{R}^{m \times r}$  tiene sus columnas linealmente independiente, y  $N \in \mathbb{R}^{m \times (n-r)}$ . Calculamos

$$\nabla C_c^T s = [B|N] \begin{pmatrix} s_B \\ s_N \end{pmatrix} = B s_B + N s_N = 0$$

$$B s_B = -N s_N. \quad (6.17)$$

- Si  $\text{rang}(\nabla C_c^T) = m$ , entonces de (6.17) resulta

$$s_B = -B^{-1} N s_N. \quad (6.18)$$

- Si  $\text{rang}(\nabla C_c^T) = r \leq m$ , entonces como la matriz  $B^T B$  es no singular

$$\begin{aligned} B^T B s_B &= -B^T N s_N \\ s_B &= -(B^T B)^{-1} B^T N s_N \\ &= -B^+ N s_N \end{aligned} \quad (6.19)$$

donde  $B^+$  es la inversa a izquierda o matriz pseudoinversa de  $B$ .

Consideremos el caso que  $s_B$  está dado por (6.18).

$$\begin{aligned} s &= \begin{pmatrix} s_B \\ s_N \end{pmatrix} = \begin{pmatrix} -B^{-1}N s_N \\ s_N \end{pmatrix} = \begin{pmatrix} -B^{-1}N \\ I_{n-m} \end{pmatrix} s_N \\ &= W s_N, \end{aligned}$$

donde

$$W = \begin{pmatrix} -B^{-1}N \\ I_{n-m} \end{pmatrix} \in \mathbb{R}^{n \times (n-m)}$$

matriz que llamamos **reducción en**  $\mathcal{N}(\nabla C_c^T)$ .

Analizamos la función cuadrática,

$$\begin{aligned} q_c(s) &= q_c(s_B, s_N) = q(W s_N, s_N) = \bar{q}_c(s_N) \\ &= \frac{1}{2}(W s_N)^T H_c(W s_N) + \nabla \ell_c^T W s_N \\ &= \frac{1}{2}s_N^T W^T H_c W s_N + \nabla \ell_c^T W s_N. \end{aligned}$$

Definimos

$$\textbf{Hessiana reducida: } W^T H_c W \in \mathbb{R}^{(n-m) \times (n-m)}.$$

$$\textbf{Gradiente reducido: } W^T \nabla \ell_c \in \mathbb{R}^{(n-m)}.$$

**Comentario:** En la implementación nunca se forma las matrices  $W$  y  $W^T H_c W$ .

Consideremos una partición de  $H_c$  de la siguiente manera

$$H = \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^T & H_{22} \end{pmatrix}$$

$H_{11} \in \mathbb{R}^{m \times n}$ ,  $H_{12} \in \mathbb{R}^{m \times (n-m)}$ ,  $H_{22} \in \mathbb{R}^{(n-m) \times (n-m)}$ . Así, la matriz Hessiana reducida se puede escribir como

$$\begin{aligned} W^T H_c W &= (-(B^{-1}N)^T \mid I) \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^T & H_{22} \end{pmatrix} \begin{pmatrix} -B^{-1}N \\ I_{n-m} \end{pmatrix} \\ &= (-(B^{-1}N)^T H_{11} + H_{12}^T \mid -(B^{-1}N)^T H_{12} + H_{22}) \begin{pmatrix} -B^{-1}N \\ I_{n-m} \end{pmatrix} \\ &= (B^{-1}N)^T H_{11} B^{-1}N - H_{12}^T B^{-1}N - (B^{-1}N)^T H_{12} + H_{22} \\ &= (B^{-1}N)^T H_{11} B^{-1}N - ((B^{-1}N)^T H_{12})^T - (B^{-1}N)^T H_{12} + H_{22}. \end{aligned} \quad (6.20)$$

El vector gradiente reducido se puede escribir

$$\begin{aligned} W^T \nabla \ell_c = W^T g &= W^T \begin{pmatrix} g_B \\ g_N \end{pmatrix} \\ &= -B^{-1} N g_B + g_N \end{aligned} \quad (6.21)$$

La siguiente proposición establece propiedades de la matriz  $W$ .

**Proposición 6.1**

- a) Las columnas de  $W$  forman una base de  $\mathcal{N}(\nabla C_c^T)$ .
- b)  $\text{Col}(W) = \mathcal{N}(\nabla C_c^T)$
- c)  $W(W^T W)^{-1} W^T$  es la proyección  $\ell_2$  en  $\mathcal{N}(\nabla C_c^T)$

**Demostración:** Ejercicio. ■

Usando esta estrategia, basada en la técnica de gradiente reducido, hemos transformado el problema cuadrático de dimensión  $n$  en un problema sin restricciones, restringido al espacio tangente de las restricciones. La dimensión del problema es  $n - r$ , donde  $r = \text{rang}(\nabla C_c^T)$ ,

$$\min \bar{q}_c(s_N) = \frac{1}{2} s_N^T W^T H_c W s_N + \nabla \ell_c^T W s_N.$$

Para resolver este problema podemos usar un método Newton inexacto o direcciones conjugadas.

A este punto se dispone de un método para resolver

$$\begin{cases} \min & q_c(s) = \frac{1}{2} s^T H_c s + \nabla \ell_c^T s \\ \text{s.a} & \\ & \nabla C_c^T s = 0. \end{cases}$$

Veamos ahora como se procede cuando  $C(x_c) \neq 0$ .

Supongamos que se dispone de un punto linealmente factible,  $s_{LF}$ . Entonces,

$$\begin{aligned} \nabla C_c^T s &= C_c \Rightarrow \nabla C_c^T s = \nabla C_c^T s_{LF} \\ \nabla C_c^T (s - s_{LF}) &= 0. \end{aligned} \quad (6.22)$$

Hacemos un cambio de variable

$$S = s - s_{LF} \Rightarrow s = S + s_{LF}$$

El sistema  $\nabla C_c^T(s - s_{LF}) = 0$  se transforma en  $\nabla C_c^T S = 0$ .

$$\begin{aligned}
 q(s) = q(S + s_{LF}) &= \frac{1}{2}(S + s_{LF})^T H_c(S + s_{LF}) + \nabla \ell_c^T(S + s_{LF}) \\
 &= \frac{1}{2}S^T H_c S + S^T H_c s_{LF} + \frac{1}{2}s_{LF}^T H_c s_{LF} + \nabla \ell_c^T S + \nabla \ell_c^T s_{LF} \\
 &= \frac{1}{2}S^T H_c S + (H_c s_{LF} + \nabla \ell_c)^T S + q_c(s_{LF}) \quad (6.23) \\
 &= Q(S). \quad (6.24)
 \end{aligned}$$

Entonces, dado el problema

$$(PC) \equiv \begin{cases} \min & q_c(s) = \frac{1}{2}s^T H_c s + \nabla \ell_c^T s \\ \text{s.a} & \\ & \nabla C_c^T s + C_c = 0. \end{cases}$$

- Se busca un punto linealmente factible  $s_{LF}$  y se transforma el problema en

$$(PCT) \equiv \begin{cases} \min & q_c(S + s_{LF}) = \frac{1}{2}S^T H_c S + (H_c s_{LF} + \nabla \ell_c)^T S + q_c(s_{LF}) \\ \text{s.a} & \\ & \nabla C_c^T S = 0. \end{cases}$$

- Se hace la reducción al espacio tangente de las restricciones

$$(PCTR) \equiv \min \bar{q}_c(S_N) = \frac{1}{2}S_N^T W^T H_c W S_N + \nabla \ell_c W S_N.$$

Sea  $S_N^* \in \mathbb{R}^{n-r}$  la solución de este problema.

- Usando las transformaciones

$$\begin{aligned}
 S_c &= W s_N^* = \begin{pmatrix} -B^+ N S_N^* \\ S_N^* \end{pmatrix} \\
 s_c &= S_c + s_{LF}
 \end{aligned}$$

Queda por ver como se halla el punto  $s_{LF}$ , solución, podría ser en el sentido de los cuadrados mínimos, de  $\nabla C_c^T s + C_c = 0$ .

- Si  $\text{rang}(\nabla C_c^T) = m \ll n$ , usar el método de Craig para resolver el sistema lineal  $\nabla C_c^T s = -C_c$ , haciendo el cambio de variables  $s = \nabla C_c y$  con  $y \in \mathbb{R}^m$ . Como  $\nabla C_c$  es de rango completo, la matriz  $\nabla C_c^T \nabla C_c$  es definida positiva y el sistema

$$\nabla C_c^T \nabla C_c y = -C_c$$

se puede resolver usando algún método de direcciones conjugadas. Sea  $y_*$  la solución. Determinamos una de las soluciones de  $\nabla C_c^T s + C_c = 0$  mediante

$$\begin{aligned} y_* &= -(\nabla C_c^T \nabla C_c)^{-1} C_c \\ \nabla C y_* &= -\nabla C_c (\nabla C_c^T \nabla C_c)^{-1} C_c \\ s_{LF} &= -\nabla C_c (\nabla C_c^T \nabla C_c)^{-1} C_c. \end{aligned} \quad (6.25)$$

Es claro que la solución (6.25) es la solución de norma  $\ell_2$ - mínima.

- Si  $\text{rang}(\nabla C_c^T) = m \approx n$ , hallar  $s_{LF}$  como solución aproximada de

$$(PNM) \equiv \begin{cases} \min & \|s\|_2^2 \\ \text{s.a} & \\ & \nabla C_c^T s + C_c = 0 \end{cases}$$

usando reducción en  $\mathcal{N}(\nabla C_c^T)$  y direcciones conjugadas. Notemos que la matriz Hessiana de la función cuadrática es la identidad.

- Si  $\text{rang}(\nabla C_c^T) = r < m$ , hallar  $s_{LM} = s_{CM}$  como solución aproximada de

$$\min \frac{1}{2} \|\nabla C_c^T s + C_c\|_2^2 = \frac{1}{2} s^T \nabla C_c \nabla C_c^T s + (\nabla C_c C)^T s + \|C_c\|_2^2$$

usando direcciones conjugadas.

**Algoritmo 6.3 (Método de Dennis-Maciel para resolver el subproblema cuadrático)**

**I) Factibilidad**

1. Hallar  $B$  y  $N$  tal que  $\nabla C_c^T = [B|N]$
2. Calcular  $s_{LF}$ 
  - Si  $\text{rang}(\nabla C_c^T) = m \ll n$ , aplicar el método de Craig.
  - Si  $\text{rang}(\nabla C_c^T) = m \approx n$ , hallar  $s_{LF}$  como solución de

$$(PNM) \equiv \begin{cases} \min & \|s\|_2^2 \\ \text{s.a} & \nabla C_c^T s + C_c = 0 \end{cases}$$

- Si  $\text{rang}(\nabla C_c^T) = r < m$ , hallar  $s_{LM}$  como solución aproximada de

$$\min \frac{1}{2} \|\nabla C_c^T s + C_c\|_2^2.$$

**II) Optimalidad**

Transformar el problema cuadrático en

$$(PCT) \equiv \begin{cases} \min & q_c(s) \\ \text{s.a} & \nabla C_c^T (s - s_{LF}) = 0 \end{cases}$$

y luego hacer la reducción al espacio tangente de las restricciones.

**Comentarios:** Dado que se propone usar direcciones conjugadas, algoritmo 6.3 detecta la existencia de direcciones que yacen en el espacio tangente que son de curvatura negativa o cero para la matriz Hessiana reducida  $W^T H_c W$  y por lo tanto para  $H$ .

### 6.2.6 Ejercicios

**Ejercicio 6.3** Demuestre proposición 6.1.

# Referencias

- [1] J. ABADIE and J. CARPENTER. Generalization of the Wolfe reduced gradient method to the case of nonlinear constraints. In R. Fletcher, editor, *Optimization*. Academic Press, New York, 1969.
- [2] M.S. BAZARAA, H.D. SHERALI, and C.M. SHETTY. *Nonlinear Programming, Theory and Algorithms*. Wiley & Sons, New York, 1993. Second edition.
- [3] D.P. BERTSEKAS. *Nonlinear Programming*. Athenas Scientific, Belmont, MA, 1999. Second Edition.
- [4] F.H. CLARKE. *Optimization and Nonsmooth Analysis*. SIAM, Philadelphia, Pennsylvania, 1990.
- [5] T.F. COLEMAN. Large sparse numerical optimization. Springer-Verlag, New York, Heidelberg, Berlin, 1984.
- [6] G.B. DANTZIG and M. THAPA. *Linear Programming 1: Introduction*. Springer-Verlag, New York, 1997.
- [7] R.S. DEMBO, S.C. EISENSTAT, and T. STEIHAUG. Inexact Newton methods. *SIAM Journal on Numerical Analysis*, 19(2):400–408, 1982.
- [8] R.S. DEMBO and T. STEIHAUG. Truncated-Newton algorithms for large-scale unconstrained optimization. *Mathematical Programming*, 26:190–212, 1983.
- [9] J.E. DENNIS and M.C. MACIEL. A conjugate direction algorithm for equality constrained quadratic programming, 1988. Reporte de circulación local en Rice University.
- [10] J.E. DENNIS and J.J. MORE. A characterization of superlinear convergence and applications to quasi-Newton methods. *Mathematics of Computation*, 28(3126):549–560, 1974.



- [11] J.E. DENNIS and J.J. MORE. Quasi-Newton methods, motivation and theory. *SIAM Review*, 19(1):46–89, 1977.
- [12] J.E. DENNIS and R.B. SCHNABEL. *Numerical methods for unconstrained optimization and nonlinear systems*. Prentice-Hall, Englewood Cliffs, New Jersey, 1983.
- [13] A.V. FIACCO and G.P. McCORMICK. *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*. SIAM, Philadelphia, Pennsylvania, 1990.
- [14] R. FLETCHER. *Practical methods of optimization*. J.Wiley & Sons, New York, 1987.
- [15] F. GOMES, M.C.MACIEL, and J.M. MARTINEZ. Nonlinear programming algorithms using trust regions and augmented Lagrangians with nonmonotone penalty parameters optimization. *Mathematical Programming*, 84:161–200, 1999.
- [16] J-B. HIRIART-URRUTY and C. LEMARECHAL. *Convex Analysis and Minimization Algorithms I*. Springer-Verlag, Berlin New York, 1993.
- [17] C.T. KELLEY. *Iterative Methods for Linear and Nonlinear Equations*. SIAM, Philadelphia, Pennsylvania, 1995.
- [18] H.W. KUHN. Nonlinear Programming: A historical view. In R.W. COTTLE and C.E. LEMKE, editors, *Nonlinear Programming*, volume IX, pages 1–26. American Mathematical Society, 1976.
- [19] L.S. LASDON. Reduced gradient methods. In M.J.D. POWELL, editor, *Optimization 81*, pages 243–250. Academic Press, London, 1981.
- [20] D. LUENBERGER. *Linear and Nonlinear Programming*. Addison Wesley, New York, 1989. Second edition.
- [21] M.C. MACIEL. *A global convergence theory for a general class of trust region algorithms for equality constrained optimization*. PhD thesis, Department of Mathematical Science, Rice University, Houston, Texas, 1992.
- [22] M.C. MACIEL. Notas de clase: Optimización con restricciones. 8000 Bahía Blanca, 2000. Dep. de Matemática, Universidad Nacional del Sur.
- [23] G.P. McCORMICK. *Nonlinear Programming: Theory, Algorithms and Applications*. Wiley & Sons, New York, 1983.

- [24] J.J. MORE. A collection of nonlinear model problems. In *Lectures in Applied Mathematics*, volume 26. American Mathematical Society, Providence, Rhode Island, 1990. Also available as MCS-P60-0289, (1989).
- [25] K. MURTY. *Linear and Combinatorial Programming*. J. Wiley & Sons, New York, 1994.
- [26] G.L. NEMHAUSER and L.A. WOLSEY. *Integer and Combinatorial Optimization*. Wiley-Interscience, New York, 1988.
- [27] J.M. ORTEGA and W.C. RHEINBOLDT. *Iterative Solution of Nonlinear Equations in Several Variables*. Academic Press, New York, 1970.
- [28] M. PADBERG. *Linear Optimization and Extensions*. Springer-verlag, New York, 1997.
- [29] W.C. RHEINBOLDT. *Methods for Solving Systems of Nonlinear Equations*. SIAM, Philadelphia, Pennsylvania, 1998.
- [30] R.T. ROCKAFELLAR. *Convex Analysis*. Princeton University Press, Princeton, 1970.
- [31] J.B. ROSEN. The gradient projection method for nonlinear programming. Part I. linear constraints. *Journal Soc. Indust. Appl. Math.*, 8(1):181–217, 1960.
- [32] J.B. ROSEN. The gradient projection method for nonlinear programming. Part II. nonlinear constraints. *Journal Soc. Indust. Appl. Math.*, 9(4):514–532, 1961.
- [33] R.W.H. SARGENT. Reduced-gradient and projection methods for nonlinear programming. In W. Murray, editor, *Numerical methods for constrained optimization*, pages 149–174. Academic Press, London, 1974.
- [34] R.W.H. SARGENT and B.A. MURTAGH. Projection methods for non-linear programming. *Mathematical Programming*, 4:245–268, 1973.
- [35] M. SIMONNARD. *Programación Lineal*. Paraninfo, Madrid, 1972.
- [36] R.A. TAPIA. Lecture notes: Nonlinear programming. Houston, Texas 77251, 1988. Rice University.
- [37] K.A. WILLIAMSON. *A robust trust region algorithm for nonlinear programming*. PhD thesis, Department of Mathematical Science, Rice University, Houston, Texas, 1990.