

# Algoritmos para la implementación computacional del método de Gauss-Newton para el problema de aproximación no lineal

Alessandri Canchoa Q. <sup>1</sup>

## Resumen

En el presente trabajo de investigación se desarrollan los algoritmos para la implementación computacional del método de Gauss-Newton para el problema de aproximación no lineal, en particular para el problema de mínimos cuadrados no lineal.

**Palabras clave:**

## Abstract

Presently investigation work the algorithms are developed for the implementation computacional of Gauss-Newton method for the problem of non lineal approach, in particular for the nonlinear least-square problems.

**Key words:**

## 1. Introducción

En el presente trabajo se enfoca el problema de la aproximación no lineal aplicando el método de Gauss-Newton. Una de las aplicaciones se da con la norma euclídeana para solucionar el problema de aproximación no lineal conocido como regresión no lineal, ajuste no lineal o problema de mínimos cuadrados no lineal. Cuando se requiere ajustar los datos obtenidos a una curva, generalmente se trata de transformar los datos para tener un problema de ajuste lineal, esto no siempre es posible y en caso de serlo se introduce un error; pueden variar sustancialmente las soluciones obtenidos del problema original y del problema transformado. En estos casos es recomendable realizar un ajuste no lineal en caso de conocerse el modelo y en otro caso el investigador podría proponer un nuevo modelo.

El objetivo del presente trabajo es el desarrollo de algoritmos para la implementación computacional del método de Gauss-Newton para el problema de aproximación no lineal, para lo cual se desarrollan los fundamentos teóricos. Este problema de aproximación no lineal es un problema de minimización sin restricciones con una función objetivo muy especial. En el método de Gauss-Newton en cada paso se requiere encontrar una dirección descendente, esto se obtiene solucionando un problema de aproximación lineal, para lo cual se puede aplicar la factorización QR mediante las desviaciones de Householder. Se desarrolla el algoritmo para solucionar el problema de aproximación lineal mediante la factorización QR considerando el evitar el problema de desbordamientos por exceso o por defecto (overflow y underflow). También se propone un criterio para analizar que es muy importante en todo algoritmo. También se muestra la manera de aplicar estos algoritmos y como en el caso de aproximación lineal también se obtiene el coeficiente de determinación múltiple R<sup>2</sup> y análisis de varianza para probar la

significación de la regresión.

## 2. Revisión de literatura

En este capítulo se presentan los conceptos básicos, definiciones, notaciones y algunos resultados que serán requeridos posteriormente.

### 1.1. La factorización QR

Este tipo de factorización es aplicado en (Houch, 1997) para solucionar el problema de mínimos cuadrados con pesos.

**Definición 1.1** Sea  $A \in \mathbb{R}^{m \times n}$ ,  $m \geq n$ . Se dice que la matriz A admite la factorización QR, si existe una matriz ortogonal  $Q \in \mathbb{R}^{m \times n}$  ( $Q^T Q = I$ ) y una matriz  $R = \begin{pmatrix} R_1 \\ 0 \end{pmatrix} \in \mathbb{R}^{m \times n}$  con una matriz triangular superior tal que  $A = QR$ .

### Teorema 1.2

Sea  $A \in \mathbb{R}^{m \times n}$ ,  $\text{rang}(A) = n$ ,  $m \geq n$ . Entonces A tiene una factorización QR donde  $R = \begin{pmatrix} R_1 \\ 0 \end{pmatrix} \in \mathbb{R}^{m \times n}$  con  $R_1 \in \mathbb{R}^{n \times n}$  una matriz superior con elementos diagonales distintos de cero.

**Prueba:** Sean  $a_1, \dots, a_n$  las columnas linealmente independientes de A. Respecto a estas columnas se aplica el método de la ortonormalización de E. Schmidt. Resultando un sistema ortonormal  $\{a_1, \dots, a_n\} \in \mathbb{R}^m$  con  $\text{Lin}\{a_1, \dots, a_k\} = \text{Lin}\{q_1, \dots, q_k\}$ ,  $k = 1, \dots, n$ . En especial se cumple

$$a_k = \sum_{j=1}^k r_{jk} q_j, \quad k = 1, \dots, n$$

y  $r_{kk} \neq 0$ . Completando  $\{q_1, \dots, q_n\}$  a un sistema ortonormal  $\{q_1, \dots, q_n, q_{n+1}, \dots, q_m\} \in \mathbb{R}^m$ , se obtiene  $Q := (q_1, \dots, q_m) \in \mathbb{R}^{m \times m}$  es ortogonal. Definiendo

<sup>1</sup> Profesor Principal, Facultad de Ciencias, Departamento de Matemática de la UNALM. Lima, Perú. E-mail: canchoa@yahoo.com.

$$(R_1)_{jk} := \begin{cases} r_{jk} & \text{para } j \leq k \\ 0 & \text{para } j > k \end{cases}$$

se cumple  $A = (a_1, \dots, a_n) = QR$ ,  $R := \begin{pmatrix} R_1 \\ 0 \end{pmatrix}$

### 1.2. Factorización QR de Householder

Dada la matriz  $A \in \mathbb{R}^{m \times n}$ , para los métodos numéricos para calcular los valores propios, así como para la solución de sistemas lineales solamente es importante el caso especial  $m \neq n$ . La utilización de la factorización QR para solucionar problemas de la aproximación lineal exige también el caso  $m \neq n$ .

La idea del método de Householder es el siguiente. En el primer paso se determina una matriz ortogonal  $P_1 \in \mathbb{R}^{m \times m}$  tal que

$$P_1 A = \left( \begin{array}{c|ccc} * & * & \dots & * \\ \hline 0 & * & \dots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \dots & * \end{array} \right)$$

Al multiplicar por  $P_1$  la primera columna de  $A$  se vuelve un múltiplo del primer vector canónico.

Supongamos que fueron determina dos las matrices ortogonales  $P_1, \dots, P_k$  tal que  $P_k \dots P_1 A$  tenga la siguiente forma

$$P_k \dots P_1 A = \left( \begin{array}{c|ccc} * & * & \dots & * \\ \hline 0 & * & \dots & * \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & * \\ \hline 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{array} \right) \left. \begin{array}{l} k \\ m-k \end{array} \right\}$$

$$= \left( \begin{array}{c|c} R^k & A_{12}^k \\ \hline 0 & A_{22}^k \end{array} \right)$$

newtheorem {ejemplo} [teorema] {Ejemplo} donde  $R^k \in \mathbb{R}^{k \times k}$  es una matriz triangular superior. Sea  $k \leq n-1$ , de otro modo  $P_k \dots P_1 A$  tendría la forma deseada. Entonces se determina una matriz ortogonal  $\tilde{P}_{k+1} \in \mathbb{R}^{(m-k) \times (m-k)}$ , tal que la primera

columna de  $\tilde{P}_{k+1} A_{22}^k$  sea un múltiplo del primer vector canónico de  $\mathbb{R}^{m-k}$ .

Se define

$$P_{k+1} := \left( \begin{array}{c|c} \mathbb{I}_k & 0 \\ \hline 0 & \tilde{P}_{k+1} \end{array} \right),$$

donde  $R^{k+1}$  es una matriz identidad. Entonces  $P_{k+1}^k \mathbb{R}^{m \times m}$  es evidentemente ortogonal. Además se cumple

$$P_{k+1} P_k \dots P_1 A = \left( \begin{array}{c|c} R^k & A_{12}^k \\ \hline 0 & \tilde{P}_{k+1} A_{22}^k \end{array} \right)$$

$$= : \left( \begin{array}{c|c} R^{k+1} & A_{12}^{k+1} \\ \hline 0 & A_{22}^{k+1} \end{array} \right)$$

donde  $R^{k+1}$  es una matriz triangular superior. Es importante que las primeras  $k$  las no cambien al multiplicar a  $P_k \dots P_1 A$  por la izquierda con  $P_{k+1}$ .

**Definición 1.3** Se dice que  $P \in \mathbb{R}^{n \times n}$ , es una matriz de Householder, si es de la forma

$$P = I - \frac{2uu^T}{u^T u}$$

donde  $u \in \mathbb{R}^n \setminus \{0\}$ .

Las matrices de Householder son simétricas y ortogonales.

Definimos

$$\text{sign}(t) = \begin{cases} 1 & , \text{ si } t \geq 0 \\ -1 & , \text{ si } t < 0 \end{cases}$$

**Lema 1.4** Para  $u \in \mathbb{R}^n \setminus \{0\}$  la matriz de Householder

$$P := I - \frac{2uu^T}{u^T u} = I - \beta uu^T$$

con

$$u := x + \text{sign}(x_1) \|x\|_2 e_1 \quad \beta = \frac{2}{u^T u} = \frac{1}{\|x\|_2 (\|x\|_2 + |x_1|)}$$

$$Px = -\text{sign}(x_1) \|x\|_2 e_1 = (*, 0, 0, \dots, 0)^T.$$

, se cumple

**Prueba.**

$$\beta = \frac{2}{u^T u}$$

$$= \frac{2}{(x + \text{sign}(x_1) \|x\|_2 e_1)^T (x + \text{sign}(x_1) \|x\|_2 e_1)}$$

$$= \frac{2}{\|x\|_2^2 + 2\|x\|_2 |x_1| + \|x_2\|^2}$$

$$= \frac{1}{\|x\|_2 (\|x\|_2 + |x_1|)}$$

$$Px = Pu - \text{sign}(x_1) \|x\|_2 P e_1$$

$$= -u - \text{sign}(x_1) \|x\|_2 e_1 + \beta u_1 \text{sign}(x_1) \|x\|_2 u$$

$$= -u - \text{sign}(x_1) \|x\|_2 e_1 + \beta (x_1 + \text{sign}(x_1) \|x\|_2) \|x\|_2 \text{sign}(x_1) u$$

$$= -u - \text{sign}(x_1) \|x\|_2 e_1 + \frac{(|x_1| + \|x\|_2) \|x\|_2}{\|x\|_2 (|x_1| + \|x\|_2)} u$$

$$= -\text{sign}(x_1) \|x\|_2 e_1$$

Se obtiene el siguiente algoritmo:

Dada la matriz

$$A = (a_{ij}) \in \mathbb{R}^{m \times n}, m \geq n$$

Para  $k = 1, 2, \dots, \min\{n, m-1\}$

Calcular una matriz  $(m-k+1) \times (m-k+1)$  de Householder con

$$\tilde{P}_{k+1} (r_{kk}, \dots, r_{mk})^T = (\alpha_k, 0, \dots, 0)^T$$

Hacer  $P_k := \text{diag}(\mathbb{I}_{k-1}, \tilde{P}_{k+1})$

Calcular  $R = P_k R$

Salida  $R = \begin{bmatrix} R_1 \\ 0 \end{bmatrix} \in \mathbb{R}^{m \times n}$  con  $R_1 \in \mathbb{R}^{n \times n}$  una matriz triangular superior.

Observaciones:

a) Definiendo  $Q := P_1 \dots P_n$  se encuentra la factorización QR dado que  $Q^T A = P_n \dots P_1 A = R$ . Para obtener la matriz R, no es necesario almacenar la matriz Q.

b) La factorización QR de una matriz  $A \in \mathbb{R}^{m \times n}$  se obtiene en  $\min\{n, m-1\}$  pasos donde  $n \leq m$ . Para  $m = n$  bastan  $n-1$  pasos.

c) En el Lema (1.4), el vector  $u$ , que define la matriz de Householder, se distingue de  $x$  solamente en la primera componente. Habría sido también posible poner  $u := x - \text{sign}(x_1) \|x\|_2 e_1$ . Igualmente la matriz de Householder transpondría  $x$  en un múltiplo del vector canónico. La desventaja es que posiblemente se sustrae en la primera componente dos números casi del mismo tamaño lo que puede influir la exactitud y producir errores por redondeos.

d) En el Lema (1.4), para evitar overflow o un underflow es conveniente calcular primero

$y := x / \|x\|_\infty$  y  $P := I - \beta u u^T$  y después

Calcular $\ a_k\ _\infty := \max_{i=k, \dots, m}  a_{ik} $
$\alpha := 0$
Para $i = k, \dots, m$
$u_{ik} := a_{ik} / \ a_k\ _\infty$
$\alpha := \alpha + u_{ik}^2$
$\alpha := \sqrt{\alpha}$
$\beta_k := 1 / [\alpha +  u_{kk} ]$
$u_{kk} := u_{kk} + \text{sign}(a_{kk}) \alpha$
% Calculando $A := P_k A$ quedan las primeras $(k-1)$ filas, por lo cual.
% sólo se calcula $a_{kk}$ y $a_{ij}$ para $j = k+1, \dots, n$ , $i = k, \dots, m$ .
$a_{kk} := -\text{sign}(a_{kk}) \ a_k\ _\infty \alpha$
Para $j := k+1, \dots, n$
$s := \beta_k \sum_{i=k}^m u_{ik} a_{ij}$
Para $i := k, \dots, m$
$a_{ij} := a_{ij} - s \times u_{ik}$

### 1.3. Aproximación Lineal

Dado un conjunto de puntos  $\{(t_i, y_i)\}_{i=1}^n$  se quiere determinar los coeficientes  $a_0, a_1, \dots, a_p$  de una curva  $y = a_p t^p + \dots + a_2 t^2 + a_1 t + a_0$ , tales que

$$\min \sum_{i=1}^n (y_i - a_p t_i^p - \dots - a_2 t_i^2 - a_1 t_i - a_0)^2,$$

este problema es equivalente

$$\min \|Ax - b\|_2,$$

donde

$$A = \begin{pmatrix} t_1^p & t_1^{p-1} & \dots & t_1^2 & t_1 & 1 \\ t_2^p & t_2^{p-1} & \dots & t_2^2 & t_2 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ t_n^p & t_n^{p-1} & \dots & t_n^2 & t_n & 1 \end{pmatrix}, x = \begin{pmatrix} a_p \\ a_{p-1} \\ \vdots \\ a_1 \\ a_0 \end{pmatrix}, b = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

$$\beta = \frac{1}{\|y\|_2 (\|y\|_2 + |y_1|)}.$$

En el algoritmo para la factorización QR se tomará en cuenta el evitar desbordamiento por exceso y por defecto (overflow y underflow); existen varias propuestas dos de ellas son analizadas por Dubrulle ([5]).

**Teorema 1.5** Para cualquier matriz  $A \in \mathbb{R}^{m \times n}$ ,  $m \geq n$  existe una factorización QR.

La implementación del algoritmo sería la siguiente en el  $k$ -ésimo paso para el cálculo de

$$\tilde{P}_k = I_{m-k+1} - \beta u_k u_k^T,$$

$$u_k := (u_{kk}, \dots, u_{mk})^T, \beta_k = 2 / (u_k^T u_k)$$

tal que

$$\tilde{P}_k a_k = (r_k, 0, \dots, 0)^T \text{ con } a_k = (a_{kk}, \dots, a_{mk})^T$$

En general veces dado un conjunto de puntos  $\{(t_i, y_i)\}_{i=1}^n$  se quiere determinar los coeficientes

$a_0, a_1, \dots, a_p$  de una curva

$y = a_p \varphi_p(t) + \dots + a_2 \varphi_2(t) + a_1 \varphi_1(t) + a_0$ , que minimicen la suma de los cuadrados de las distancias de los puntos dados al polinomio.

Esto se puede expresar como sigue

$$\min \sum_{i=1}^n (y_i - a_p \varphi_p(t_i) - \dots - a_2 \varphi_2(t_i) - a_1 \varphi_1(t_i) - a_0)^2,$$

este problema es equivalente

$$\min \|Ax - b\|_2,$$

donde

$$A = \begin{pmatrix} \varphi_p(t_1) & \varphi_{p-1}(t_1) & \cdots & \varphi_2(t_1) & \varphi_1(t_1) & 1 \\ \varphi_p(t_2) & \varphi_{p-1}(t_2) & \cdots & \varphi_2(t_2) & \varphi_1(t_2) & 1 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ \varphi_p(t_n) & \varphi_{p-1}(t_n) & \cdots & \varphi_2(t_n) & \varphi_1(t_n) & 1 \end{pmatrix}, x = \begin{pmatrix} a_p \\ a_{p-1} \\ \vdots \\ a_1 \\ a_0 \end{pmatrix}, b = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

Otra aplicación es en la solución de un sistema lineal  $Ax = b$ , cuando se tienen más ecuaciones que variables, una alternativa es solucionar el problema

$$\min \|Ax - b\|, x \in \mathbb{R}^n.$$

Para las normas  $\|\cdot\|_1$  o  $\|\cdot\|_\infty$  se obtiene un problema de optimización lineal. En caso de que se trate de  $\|\cdot\|_2$ , se obtiene el problema de mínimos cuadrados.

Consideramos el problema

$$\min \|Ax - b\|_2, x \in \mathbb{R}^n, \text{ (AL)}$$

este problema es denominado un problema de aproximación lineal.

**Teorema 1.6** Sea  $A \in \mathbb{R}^{m \times n}$ ,  $m \geq n$  A 2 en el problema de la aproximación lineal (AL). Entonces

1.  $x^* \in \mathbb{R}^n$  es solución de (AL), si y sólo si  $A^T Ax^* = A^T b$ , es decir se cumplen las ecuaciones normales.

2. El conjunto  $L := \{x \in \mathbb{R}^n : A^T Ax = A^T b\}$  de las soluciones de (AL) es un subespacio afín no vacío

$$\begin{aligned} f(x) - f(x^*) &= \frac{1}{2}((x^T A^T Ax - 2x^T A^T b + \|b\|_2^2) - \frac{1}{2}((x^{*T} A^T Ax^* - 2x^{*T} A^T b + \|b\|_2^2)) \\ &= \frac{1}{2}((x^T A^T Ax - 2x^T A^T b) - \frac{1}{2}((x^{*T} A^T Ax^* - 2x^{*T} A^T b)) \\ &= \frac{1}{2}((x^T A^T Ax - 2x^T A^T b) - \frac{1}{2}((x^{*T} A^T b - 2x^{*T} A^T b)) \\ &= \frac{1}{2}((x^T A^T Ax - 2x^T A^T b + x^{*T} A^T b)) \\ &= \frac{1}{2}((x^T A^T Ax - 2x^{*T} A^T Ax + x^{*T} A^T Ax^*)) \\ &= \frac{1}{2}(x - x^*)^T A^T A(x - x^*) \\ &= \frac{1}{2}\|A(x - x^*)\|_2^2 \\ &\geq 0 \end{aligned}$$

lo cual implica que  $f(x) \geq f(x^*)$ , entonces  $x^*$  es una solución de AL).

2. Hay que mostrar que  $L \neq \emptyset$ .

Se cumple: El sistema  $Ax = b$  es solucionable si sólo si

$$y \in \text{Nucleo}(A^T) \text{ implica } b^T y = 0.$$

En efecto:  $\implies Ax = b$  tiene solución

Lo que implica

de  $\mathbb{R}^n$ , es decir con  $x_1, x_2 \in L$  tales  $(1 - \lambda)x_1 + \lambda x_2 \in L, \forall \lambda \in \mathbb{R}$ . Además se cumple  $Ax_1 = Ax_2, \forall x_1, x_2 \in L$ .

3. (AL) tiene una solución única si y sólo si  $\text{Rang}(A) = n$ .

4. De todas las soluciones de (AL) existe exactamente una solución  $x_{AL}$  con norma de Euclides minimal.

**Prueba. 1.** Se define  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  por

$$\begin{aligned} f(x) &= \frac{1}{2}\|Ax - b\|_2^2 \\ &= \frac{1}{2}(Ax - b)^T (Ax - b) \end{aligned}$$

solucionar el problema (AL) es equivalente a solucionar  $\min f(x)$ .

$\implies$  Sea  $x^* \in \mathbb{R}^n$  una solución de (AL). Entonces

$$\nabla f(x^*) = 0 \implies A^T Ax^* - A^T b = 0$$

esto es,  $A^T Ax^* = A^T b$ . Luego  $x^*$  es una solución de las ecuaciones normales.

$\Leftarrow$  Sea  $x^*$  una solución de las ecuaciones normales, entonces

$$y^T Ax = (A^T y)^T x = y^T b, y \in \mathbb{R}^m.$$

$$y \in \text{Nucleo}(A^T) \implies A^T y = 0 \implies y^T b = 0$$

$$\Leftarrow y \in \text{Nucleo}(A^T) \implies y^T b = 0.$$

Sean  $x; b$  linealmente dependientes

$$\implies x \in \text{Nucleo}(A^T) \implies \text{Rg}(A^T) = \text{Rg}(A) = \text{Rg}(A | b).$$

Sea  $x \in \text{Núcleo}$

$(A^T A)^T = \text{Núcleo}(A^T A)$ . De esto se

obtiene  $A^T A x = 0$  y se cumple

$$0 = x^T A^T A x = \|Ax\|_2^2,$$

de donde se sigue que  $Ax = 0$  y

$$x^T A^T b = (Ax)^T b = 0,$$

de esto se obtiene que  $x \in L$ , es decir,  $L \neq \emptyset$ .

$L$  es subespacio ajeno  $\iff$

$$\begin{aligned} A^T A((1-\lambda)x_1 + \lambda x_2) &= (1-\lambda)A^T A x_1 + \lambda A^T A x_2 \\ &= (1-\lambda)^T b + \lambda A^T b \\ &= A^T b \end{aligned}$$

Sea  $x_1, x_2 \in L \Rightarrow f(x_1) = f(x_2)$  Como

$$0 = f(x_1) - f(x_2) = \frac{1}{2} \|A(x_1 - x_2)\|_2^2,$$

se obtiene  $Ax_1 = Ax_2$ .

3. Núcleo  $Ax_1 = Ax_2$ .

Entonces  $A^T A$  es regular, si  $\text{Rg}(A) = n$ .

4. Existencia: hay que demostrar que el problema

$\min \|x\|_2$  sujeto a  $x \in L := \{x \in \mathbb{R}^n : A^T A x = A^T b\}$  tiene una solución única.

Sea  $x^* \in L$ , luego el conjunto  $L \cap \{x \in \mathbb{R}^n : \|x\|_2 \leq \|x^*\|_2\} =: B$  es

compacto por ser la intersección de un conjunto cerrado y el otro compacto. Entonces la función  $\|\cdot\|_2$  continua tiene un mínimo en  $B \subseteq L$ , por consiguiente también toma su mínimo sobre  $L$ .

Supongamos que  $x_1, x_2 \in L$  sean dos soluciones con norma minimal, entonces  $\|x_1\|_2 = \|x_2\|_2$ .

Como  $L$  es afín  $\frac{1}{2}(x_1 + x_2) \in L$  y

$$\begin{aligned} \left\| \frac{1}{2}(x_1 + x_2) \right\|_2 &\leq \frac{1}{2}(\|x_1\|_2 + \|x_2\|_2) \\ &= \|x_1\|_2 = \|x_2\|_2 \\ &\leq \left\| \frac{1}{2}(x_1 + x_2) \right\|_2 \end{aligned}$$

De esto se obtiene

$$\frac{1}{2} \|x_1 + x_2\|_2 = \|x_1\|_2$$

$$\begin{aligned} \|x_1 - x_2\|_2^2 &= \|x_1\|_2^2 - 2x_1^T x_2 + \|x_2\|_2^2 \\ &= \|x_1\|_2^2 - (\|x_1 + x_2\|_2^2 - \|x_1\|_2^2 - \|x_2\|_2^2) + \|x_2\|_2^2 \\ &= 2(\|x_1\|_2^2 + \|x_2\|_2^2) - \|x_1 + x_2\|_2^2 \\ &= 0, \end{aligned}$$

Numéricamente, el problema lineal (AL) se trata de la manera siguiente:

Por el método de Householder se determina una matriz ortogonal  $Q \in \mathbb{R}^{m \times m}$  tal que

$$Q^T A = R = \begin{pmatrix} R_1 \\ 0 \end{pmatrix} \begin{matrix} \}n \\ \}m-n \end{matrix} \quad \text{y} \quad Q^T b = \begin{pmatrix} c \\ d \end{pmatrix} \begin{matrix} \}n \\ \}m-n \end{matrix}$$

donde  $R_1 \in \mathbb{R}^{n \times n}$  una matriz triangular superior.

Si  $\text{rang}(A) = n$ ,  $R_1$  es regular. Luego los elementos diagonales son distintos de cero.

Además, si  $x \in \mathbb{R}^n$ , se cumple

$$\begin{aligned} \|Ax - b\|_2^2 &= \|Q^T (Ax - b)\|_2^2 \\ &= \|Q^T Ax - Q^T b\|_2^2 \\ &= \left\| \begin{pmatrix} R_1 \\ 0 \end{pmatrix} x - \begin{pmatrix} c \\ d \end{pmatrix} \right\|_2^2 \\ &= \|R_1 x - c\|_2^2 + \|d\|_2^2 \end{aligned}$$

La solución se obtiene de solucionar el sistema triangular  $R_1 x = c$ .

Para el caso donde la matriz  $A$  no sea regular o las columnas de  $A$  sean linealmente dependientes, hay que utilizar otros métodos. Para solucionar problemas de este tipo numéricamente se necesitan métodos para el cálculo de valores propios.

Para solucionar problemas de aproximación lineal, el método QR es más eficiente que la aplicación del método de Cholesky que soluciona las ecuaciones normales ( $A^T A x = A^T b$ ) porque el número de condición de la matriz  $A^T A$  es pésima:

$$\begin{aligned} \text{cond}(A^T A) &= \|A^T A\|_2 \|A^{-1} A^{-T}\|_2 \\ &= \rho(A^T A) \rho(A^{-1} A^{-T}) \\ &= \rho(A^2) \rho(A^2)^{-1} \\ &= \text{cond}(A^2) \end{aligned}$$

A partir de las observaciones 1.2.3 acerca del método de Householder y de estas observaciones se obtiene el siguiente algoritmo.

<b>Algoritmo QR3, para el problema de aproximación lineal:</b>
Entrada: $A = (a_{ij}) \in \mathbb{R}^{m \times n}, b = (b_i) \in \mathbb{R}^m$ donde $m \geq n$ .
Poner $Ab := [A, b]$ ; % matriz ampliada
Para $k := 1, \dots, \min(n, m - 1)$
Para $j := k : m$
$x_{j-k+1} := Ab_{j,k}$
$e_{1_{j-k+1}} := 0$
$normMax := \ x\ _\infty$
$y = x/normMax$
$alfa = \ y\ _2$
$e_{1_1} := 1$ ; % $e_1 = (1, 0, \dots, 0)^T$ en $\mathbb{R}^{(n-k)}$
Si $x_1 \geq 0$
$signx1 = 1$
sino
$signx1 = -1$
Fin-si
$u = y + signx1 * alfa * e1$
$beta = 1/(alfa * (alfa +  y_1 ))$
$Ab_{k,k} := -signx1 * normMax * alfa$
Para $i = k + 1, \dots, m$
$Ab_{(i,k)} := 0$
$V := Ab_{(k:m, (k+1):m+1)}$ % **
$A1 = Ab_{(:, 1:n)}$
$b1 = Ab_{(:, n+1)}$
Salida: $A1, b1$ donde $A1 = Q * A = \begin{bmatrix} R1 \\ 0 \end{bmatrix}$
con $R1$ triangular superior
$b1 = Q * b$

\*\* : Asigna la submatriz  $(Ab_{i,j} : i = k, \dots, m, j = k + 1, \dots, n + 1)$

### 3. Discusión

#### 3.1 Método de Gauss-Newton para problemas no lineales de aproximación

Dado un conjunto de puntos  $\{(t_i, y_i)\}_{i=1}^n$  se quiere determinar los coeficientes  $a_0, a_1, \dots, a_p$  de una curva  $y = \varphi(t; a_1, a_2, \dots, a_p)$ , tal que

$$\min \sum_{i=1}^n (y_i - \varphi(t_i; a_1, a_2, \dots, a_p))^2,$$

este problema es equivalente donde

$$\min \|H(x)\|_2, \quad H(x) = \begin{pmatrix} \varphi(t_1; x_1, x_2, \dots, x_p) - y_1 \\ \varphi(t_2; x_1, x_2, \dots, x_p) - y_2 \\ \vdots \\ \varphi(t_p; x_1, x_2, \dots, x_p) - y_p \end{pmatrix}, \quad x = \begin{pmatrix} a_1 \\ a_1 \\ \vdots \\ a_p \end{pmatrix}.$$

Consideremos el problema  $\min f(x) = \|F(x)\|, x \in \mathbb{R}^n$  (PA)

donde  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$

Asumimos que se cumplen las siguientes hipótesis

a) Dado  $x_0 \in \mathbb{R}^n$ , el conjunto de nivel  $L_0 := \{x \in \mathbb{R}^n : f(x) \leq f(x_0)\}$  es compacto.

b) Sea  $F$  continuamente diferenciable sobre un conjunto abierto que contiene a  $L_0$   
 c)  $\|F'(x) - F'(y)\| \leq \gamma \|x - y\|$  para todo  $x, y \in L_0$

**Lema 2.1** Sea  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  convexa. Entonces  $f$  es diferenciable y

$$f'(x; p) \leq f(x+p) - f(x), \text{ para todo } x, p \in \mathbb{R}^n.$$

**Prueba.** Definimos  $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}$  por

$$\Phi(t) = \frac{f(x+tp) - f(x)}{t}.$$

Mostraremos la existencia de  $\lim_{t \rightarrow 0^+} \Phi(t)$ .

Por la convexidad de  $f$  para  $t \in [0; 1]$  se cumple

$$\begin{aligned} f(x) &= f\left(\frac{1}{1+t}(x+tp) + \frac{t}{1+t}(x-p)\right) \\ &\leq \frac{1}{1+t}f(x+tp) + \frac{t}{1+t}f(x-p) \end{aligned}$$

De esto se obtiene

$$f(x) - f(x-p) \leq \frac{f(x+tp) - f(x)}{t} = \Phi(t), \forall t \in ]0, 1[$$

Esto signi.ca que  $\Phi$  esta acotada inferiormente.

Para  $0 < s \leq t < 1$  se cumple

$$\begin{aligned}\Phi(s) &= \frac{1}{s} (f(x+sp) - f(x)) \\ &= \frac{1}{s} \left( f \left( \frac{s}{t} (x+tp) + \frac{t-s}{t} x \right) - f(x) \right) \\ &\leq \frac{f(x+tp) - f(x)}{t} = \Phi(t)\end{aligned}$$

De esto se obtiene que  $\Phi$  es decreciente sobre  $]0, 1[$  y como es acotada inferiormente, existe  $f'(x; p) = \lim_{t \rightarrow 0^+} \Phi(t)$  y se cumple

$$f'(x; p) \leq f(x+p) - f(x).$$

Lema 2.2 Dado (PA) con las hipótesis a), b) y c) y con  $x_0 \in L_0$  consideremos el problema

$$\min f_x(p) = \|F(x) + F'(x)p\|, p \in \mathbb{R}^n \quad (\text{LP}_x)$$

Entonces existe una solución  $p^*$  (no necesariamente única) de  $(\text{LP}_x)$  y se cumplen.

1. Si  $f_x(p^*) = f(x)$ ,  $x$  es una solución estacionaria de (PA), es decir

$$f_x'(p; q) \geq 0 \text{ para todo } q \in \mathbb{R}^n.$$

2. Si  $f_x(p^*) < f(x)$ , para  $p \in \mathbb{R}^n$ , se cumple

$$f_x'(p; q) \leq f_x(p) - f(x) < 0,$$

es decir  $p$  es una dirección descendente para  $f$  en  $x$ . En particular,  $p^*$  es una dirección descendente para  $f$  en  $x$ , si  $f_x(p^*) \neq f(x)$ .

**Prueba.** 1) Se va a probar que  $(\text{LP}_x)$  tiene solución. Consideremos

$$\min g_x(q) := \|F(x) + q\|, q \in \text{Im}(F'(x)), \quad (\text{P}_g)$$

El conjunto  $\hat{L} := \{q \in \mathbb{R}^n : g_x(q) \leq g_x(0)\} \cap \text{Im}(F'(x))$  es cerrado por ser la intersección de conjuntos cerrados. Para  $q \in \hat{L}$  se cumple

$$\begin{aligned}\|q\| &= \|q + F(x) - F(x)\| \\ &\leq \|q + F(x)\| + \|F(x)\| \\ &\leq g_x(0) + \|F(x)\| \\ &\leq 2\|F(x)\|\end{aligned}$$

se tiene que  $\hat{L}$  también es acotado. Por consiguiente  $\hat{L}$  es compacto y el problema  $(\text{P}_g)$  tiene solución.

Una solución de  $f_x(p^*) = f(x)$  es  $p = 0$  y por lo tanto es una solución estacionaria de  $(\text{LP}_x)$ , luego se tiene

$$\begin{aligned}0 &\leq f_x'(0; q) = \lim_{t \rightarrow 0^+} \frac{f_x(0+ tq) - f_x(0)}{t} \\ &= \lim_{t \rightarrow 0^+} \frac{\|F(x) + F'(x)(tq)\| - \|F(x)\|}{t} \\ &= \lim_{t \rightarrow 0^+} \frac{\|F(x) + tF'(x)q\| - \|F(x)\|}{t} \\ &= g'(F(x); F'(x)q)\end{aligned}$$

De esto se obtiene

$$0 \leq f_x'(0; q) = g'(F(x); F'(x)q) = f'(x; q) \text{ para todo } q \in \mathbb{R}^n.$$

Por consiguiente  $x$  es una solución estacionaria de (PA).

2) Si  $f(x+p) - f(x) < 0$ , por el Lema precedente se tiene  $g'(F(x); F'(x)p) \leq g(F(x); F'(x)p) - g(F(x))$

Luego

$$\begin{aligned}f'(x; p) &= g'(F(x); F'(x)p) \\ &\leq g(F(x); F'(x)p) - g(F(x)) \\ &= f(x+p) - f(x) < 0,\end{aligned}$$

Por consiguiente  $p$  es una dirección descendente.

Una vez determinado una dirección descendente  $p$  se tiene que elegir una longitud de paso  $t$ , para lo cual se puede usar el siguiente algoritmo de Armijo.

Algoritmo: Búsqueda de Armijo
Elegir $\delta$ en $]0, 1/2[$ , y $L, u$ en $]0, 1[$ tales que $L \leq u$ ,
$\varphi_0 := 1$
Para $j := 0, 1, 2, \dots$
Si $f(x + \varphi_j p) \leq f(x) + \delta \varphi_j [f_x(p) - f(x)]$ (*)
$t := \varphi_j$ ¡Parar!
Sino
Elegir $\varphi_{j+1} \in [L \varphi_j, u \varphi_j]$
Fin-si

Este algoritmo finaliza en un número finito de pasos. En efecto, supongamos que (\*) no se cumple, luego se tendría

$$\begin{aligned}\delta [f_x(p) - f(x)] &\leq \lim_{j \rightarrow \infty} \frac{f(x + \varphi_j p) - f(x)}{\varphi_j} \\ &= f'(x; p)\end{aligned}$$

Por el Lema precedente y de esto se obtiene

$$\begin{aligned}\delta [f_x(p) - f(x)] &\leq f'(x; p) \\ &\leq f_x(x; p) - f(x)\end{aligned}$$

Finalmente se obtiene

$$0 \leq \underbrace{(1 - \delta)}_{>0} [f_x(p) - f(x)] < \underbrace{0}_{<0}$$

Lo cual es absurdo, por consiguiente (\*) se cumplirá en un número finito de pasos.

**Lema 2.3**  $x \in L_0$  y  $p \in \mathbb{R}^n$  una dirección descendente tal que

$$f_x(p) = \|F(x) + F'(x)p\| < f(x),$$

y

$$t^* = \frac{2(f_x(p) - f(x))}{\gamma \|p\|^2}.$$

Entonces

$$f(x+tp) \leq f(x) + t(f_x(p) - f(x)) + t^2 \frac{\gamma}{2} \|p\|^2$$

para todo  $t \in [0, \min(1, t^*)]$ ,

Prueba. Sea  $\hat{t}$  el primer cero positivo de  $f(x)-f(x+tp)$ . Luego

$$x + tp \in L_0 \text{ para todo } s \in [0, \hat{t}].$$

para  $t \in [0, \hat{t}]$  se cumple

$$\begin{aligned} F(x+tp) &= F(x) + tF'(x)p + \int_0^t (F'(x+sp) - F'(x))p \, ds \\ &= (1-t)F(x) + t[F(x) + F'(x)p] + \int_0^t (F'(x+sp) - F'(x))p \, ds \quad \square \end{aligned}$$

Tomando normas, para  $t \in [0, \min(1, \hat{t})]$  se obtiene

$$\begin{aligned} f(x+tp) &\leq (1-t)f(x) + tf_x(p) + t^2 \frac{\gamma}{2} \|p\|^2 \\ &= f(x) + t(f_x(p) - f(x)) + t^2 \frac{\gamma}{2} \|p\|^2 \end{aligned}$$

Para  $\hat{t} \leq 1$ , de esta inecuación con  $t = \hat{t}$  resulta que  $t^* \leq \hat{t}$  y así se obtiene  $(0, t^*) \leq \min(0, \hat{t})$

**Teorema 2.4** Dado (PA) con las hipótesis a), b) y c).

Sean  $x \in L_0$ ,  $p$  una dirección descendente. Aplicando la prueba de Armijo existe una constante tal que  $\theta > 0$

$$f(x) - f(x+tp) \geq \theta \min \left[ f(x) - f_x(p), \left( \frac{f(x) - f_x(p)}{\|p\|} \right)^2 \right].$$

**Prueba.** Sean  $\delta \in ]0, 1/2[$  y  $0 < l \leq u < 1$  y en Algoritmo de la búsqueda de Armijo con  $t = \varphi_j$ .

Para  $j = 0, t = \varphi_j = 1$  se tiene

$$f(x+p) \leq f(x) + \delta(f_x(x) - f(x))$$

Para  $j > 0$  y  $s := \varphi_{j-1}$  se cumplen dos desigualdades

$$f(x+tp) \leq f(x) + \delta t(f_x(x) - f(x)) \quad (1)$$

$$f(x+sp) > f(x) + \delta s(f_x(x) - f(x)) \quad (2)$$

Sea  $\hat{t}$  el primer cero positivo de  $f(x)-f(x+tp)$ . Se tienen dos casos

a) Para  $s \leq \hat{t}$ , por el Lema (2.3) y por (2) se obtiene

$$\begin{aligned} f(x) + \delta s(f_x(x) - f(x)) &< f(x+sp) \\ &\leq f(x) + s(f_x(p) - f(x)) + s^2 \frac{\gamma}{2} \|p\|^2 \end{aligned}$$

Simplicando se obtiene

$$0 < \frac{2l}{\gamma \|p\|^2} (\delta - 1) \leq ls \leq t$$

De esto y por (1)

$$\begin{aligned} f(x+tp) &\leq f(x) + \delta t(f_x(x) - f(x)) \\ &\leq f(x) + \frac{2l}{\gamma \|p\|^2} \delta (\delta - 1) (f_x(x) - f(x)) \end{aligned}$$

$$\frac{2l}{\gamma \|p\|^2} \delta (1 - \delta) (f(x) - f_x(p)) \leq f(x) - f(x+tp) \quad (3)$$

b) Para  $s > \hat{t}$ , por el Lema (2.3):

Como  $f(x) - f(x+\hat{t}p) = 0$ , por la prueba del Lema (2.3) se obtiene

$$\begin{aligned} 0 &\leq (f_x(p) - f(x)) + \hat{t} \frac{\gamma}{2} \|p\|^2 \\ 0 &< \frac{2}{\gamma \|p\|^2} (f(x) - f_x(p)) \leq \hat{t} < s \end{aligned}$$

Por (1) y de esto se obtiene

$$\begin{aligned} f(x+tp) &\leq f(x) + \delta t(f_x(x) - f(x)) \\ &\leq f(x) - \frac{2\delta}{\gamma \|p\|^2} (f(x) - f_x(p))^2 \end{aligned}$$

Despejando

$$f(x) - f(x+tp) \geq \frac{2\delta}{\gamma} \left( \frac{f(x) - f_x(p)}{\|p\|} \right)^2 \quad (4)$$

En (3) y (4) tomando

$$\theta = \min \left\{ \frac{2\delta}{\gamma}, \frac{2l}{\gamma \|p\|^2} \delta (1 - \delta) \right\} > 0, \text{ finalmente se obtiene}$$

$$f(x) - f(x+tp) \geq \theta \min \left[ f(x) - f_x(p), \left( \frac{f(x) - f_x(p)}{\|p\|} \right)^2 \right]$$

**Teorema 2.5** Sea  $\text{Rango}(F'(x)) = n$  para todo  $x \in L_0$ . Consideramos el siguiente algoritmo

Sea $\delta \in ]0, 1/2[$ , $0 < l \leq u < 1$ para la prueba de Armijo
$x^0 \in L_0$ % un punto inicial
Para $k := 0, 1, \dots$
Sea $p_k$ una solución de (LP <sub>k</sub> ) : minimizar $f_k(p) := \ F(x^k) + F'(x^k)p\ , p \in \mathbb{R}^n$
Si $f_k(p_k) = f(x^k)$ , entonces $x^k$ es una solución estacionaria del (PA) ¡Parar!
Sino
Calcular la búsqueda de Armijo $t_k := \varphi_j$
$x^{k+1} := x^k + t_k p_k$
Fin-si

En caso de que el algoritmo no termine en un número finito de pasos con una solución estacionaria de (PA), se tiene la sucesión  $\{x^k\} \subset L_0$  con  $f(x^k) < f(x^{k-1}), k = 0, 1, 2, \dots$  y cada punto de acumulación  $x^*$  de  $\{x^k\}$  es una solución estacionaria de (PA). Si (PA) tiene una solución estacionaria única,  $\{x^k\}$  converge a  $x^*$ .

**Prueba.** a) Sea  $x \in L_0$  y  $\|F(x) + F'(x)p\| \leq f(x)$ . Se cumple que  $\{p^k\}$  es acotado, en efecto. Se cumple

$$\begin{aligned} \|F'(x)p\| &= \|F'(x)p + F(x) - F(x)\| \\ &\leq \|F'(x)p + F(x)\| + \|F(x)\| \\ &\leq 2f(x) \end{aligned}$$

Como  $F'(x) \in \mathbb{R}^{n \times n}$  tiene rango  $n$ , se tiene que  $F'(x)^T F'(x) \in \mathbb{R}^{n \times n}$  es regular. Luego se obtiene



$$\begin{aligned}\|p\| &= \left\| [F'(x)^T F'(x)]^{-1} F'(x)^T F'(x)p \right\| \\ &\leq \left\| [F'(x)^T F'(x)]^{-1} F'(x)^T \right\| \|F'(x)p\| \\ &\leq \left\| [F'(x)^T F'(x)]^{-1} F'(x)^T \right\| (2f(x)) \\ &\leq cf(x)\end{aligned}$$

con

$$c = 2 \max_{x \in L_0} \left\| [F'(x)^T F'(x)]^{-1} F'(x)^T \right\|$$

Por consiguiente, se tiene  $\|p_k\| \leq c f(x^k) \leq c f(x^0)$ , es decir  $\{p^k\}$  es acotado.

b) Se cumple  $\lim_{k \rightarrow \infty} [f(x^k) - f_{x^k}(p_k)] = 0$ , en efecto: Por el Teorema (2.4) se tiene

$$f(x^k) - f(x^{k+1}) \geq \theta \min \left[ f(x^k) - f_{x^k}(p_k), \left( \frac{f(x^k) - f_{x^k}(p_k)}{\|p_k\|} \right)^2 \right]$$

Luego  $\lim_{k \rightarrow \infty} (f(x^k) - f(x^{k+1})) = 0$  implica  $\lim_{k \rightarrow \infty} (f(x^k) - f_{x^k}(p_k)) = 0$

c) Sea  $x^*$  un punto de acumulación de  $\{x^k\}$  y como  $\{p^k\}$  es acotado, existe un subconjunto  $K \subset \mathbb{N}$  infinito tal que  $\{x^k\}_{k \in K} \rightarrow x^*$ ,  $\{p_k\}_{k \in K} \rightarrow p^*$ . Así,  $p^*$  es una solución de

$$\min_{p \in \mathbb{R}^n} f_{x^*}(p) := \|F(x^*) + F'(x^*)p\|, \quad p \in \mathbb{R}^n \quad (\text{LP}^*)$$

con  $f_{x^*}(p^*) = f(x^*)$ . Es decir, según el Lema (2.2),  $x^*$  es una solución estacionaria.

Para  $p \in \mathbb{R}^n$  se cumple

$$\|F(x^k) + F'(x^k)p_k\| \leq \|F(x^k) + F'(x^k)p\|, \quad k = 0, 1, 2, \dots$$

Para  $k \in K$ ,  $p^*$  es solución de (LP\*). Por (b) se tiene

$$\begin{aligned}0 &= \lim_{\substack{k \rightarrow \infty \\ k \in K}} (f(x^k) - f_{x^k}(p_k)) \\ &= f(x^*) - f_{x^*}(p^*)\end{aligned}$$

Lo que implica  $f(x^*) = f_{x^*}(p^*)$ .

#### 4. Resultados

En la sección anterior se han desarrollado la base teórica del método de Gauss-Newton para problemas no lineales de aproximación para cualquier norma, también se presenta las propuestas de los correspondientes algoritmos para su implementación, en particular para la norma euclídeana  $\|\cdot\|_2$ . Para la parte estadística puede ver por ejemplo [7]. Los algoritmos correspondientes se han implementado usando el Software MATLAB, a continuación se muestran algunos ejemplos, donde se ilustrarán las corridas y los datos de entrada que se requieren en éstos métodos. Para la implementación se han considerado los siguientes criterios:

1. Para la prueba de Armijo  $\delta = 0,0001, L = 0,2500, U = 0,5000$ .

Se elige  $\varphi_{j+1} \in [L \times \varphi_j, u \times \varphi_j]$ , como el punto medio del intervalo, esto es

$$\varphi_{j+1} = 0,5 \times (L + u) \times \varphi_j.$$

2. En cada paso se tiene que solucionar el problema

(LP<sub>k</sub>) : minimizar  $f_k(p) = \|F(x^k) + F'(x^k)p\|_2, p \in \mathbb{R}^n$ , que es un problema de aproximación lineal. Para este problema se ha utilizado la factorización QR, dado que tiene la forma

$$(AL) : \min \|Az - b\|_2, z \in \mathbb{R}^n$$

donde

$$A := F'(x^k), b := -F(x^k), z := p.$$

3. Criterio de finalización: El algoritmo termina, si

$$f(x^k) - \|F(x^k) + F'(x^k)p\|_2 \leq tol$$

En los siguientes ejemplos se considera  $tol = 10^{-12}$ .

**Ejemplo 3.1** Se está buscando ajustar la función  $g(t) = ae^{bt}$  para los datos

**Tabla 3.1.1.** Datos.

$i$	1	2	3	4	5
$t_i$	0,00	1,00	2,00	3,00	4,00
$y_i$	0,60	1,90	4,30	7,60	12,6

En este caso en particular, usualmente se realiza la transformación logarítmica  $\ln y = \ln a + bt$  que tiene la forma  $Y = A + bt$  con  $Y = \ln y$  y  $A = \ln a$ . Luego, aplicando ajuste lineal y como  $a = \exp(A)$ , se obtiene

$$\hat{A} = -0,26477017804222,$$

$$\hat{\alpha} = 0,76738229101205,$$

$$\hat{b} = 0,74753392365667,$$

el coeficiente de determinación múltiple es

$$\hat{R}^2 = 0,95762278389119$$

En este ejemplo vamos a verificar que la aproximación no lineal proporciona un mejor resultado. Consideramos  $x_1 = a, x_2 = b$  de finimos la función  $F = (F_i) : \mathbb{R}^2 \rightarrow \mathbb{R}^5$  cuyas componentes están dados por

$$F_i(x) := x_1 e^{x_2 t_i} - y_i, \quad i = 1, 2, \dots, 5$$

Luego hay que solucionar el problema

$$\min f(x) := \|F(x)\|_2.$$

Se requiere

$$F'(x) = \begin{pmatrix} \nabla_x F_1(x)^T \\ \nabla_x F_2(x)^T \\ \vdots \\ \nabla_x F_5(x)^T \end{pmatrix}$$

donde

$$\nabla_x F_i(x) = \begin{pmatrix} e^{x_2 t_i} \\ x_1 t_i e^{x_2 t_i} \end{pmatrix}, \quad i = 1, 2, \dots, 5.$$

Para el punto inicial  $x_0 := (200, 30)^T$  y  $\delta = 0,0001$  en la prueba de Armijo y con  $\varepsilon = 10^{-12}$  en el criterio

para finalizar  $(f(x^k) - f_k(p_k) < \varepsilon)$ , se obtienen los

siguientes resultados

**Tabla 3.1.2** Resultados para el ejemplo 1.

k	$x_k(1)$	$x_k(2)$	$f(x_k)$	$f(x_k) - f_{x_k}(p_k)$
0	1,0000000000000000	1,0000000000000000	43,9329613297339990	43,512939070902782
1	0,85502101488523	0,84382318054493	12,7954489955334230	12,205617356995402
2	1,08340732785449	0,65357925941532	2,2784277840290081	1,382860541253122
3	1,25313196975843	0,58024245688103	0,9304863933294990	0,001628968642767
4	1,24967456136510	0,58195313522654	0,9288762008214361	0,000001679519679
5	1,25033674243286	0,58180358193354	0,9288746645852360	0,000000012260682
6	1,25028002391004	0,58181635926293	0,9288746533705558	0,000000000089237
7	1,25028487850983	0,58181526906945	0,9288746532889339	0,000000000000650

Se obtiene una aproximación de la solución óptima  $x^* \approx x^7$  y el valor óptimo  $f(x^*) \approx \|F(x^7)\|_2 = 0,9288746532889339$

con el coeficiente de determinación múltiple  $R^2=0;97108579559180$  que mejora a  $\hat{R}^2$ .  
Análisis de varianza para probar la significación de la regresión:

Así se obtiene

$$a := x_1^7 = 1,25028487850983,$$

$$b := x_2^7 = 0,58181526906945,$$

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	$F_0$
Regresión	233,12703886459491	2	116,56351943229745	67,17015498
Error o residual	6,94138755340587	4	1,73534688835147	
Total	240,06842641800077	4		

donde  $n = 5$ ;  $p = 1$ ;  $k = 2$ . Para  $\alpha = 0;05$ , como  $F_0=67;17015498 > f_{\alpha,k,n-p} = 6;94$  se rechaza la hipótesis nula.

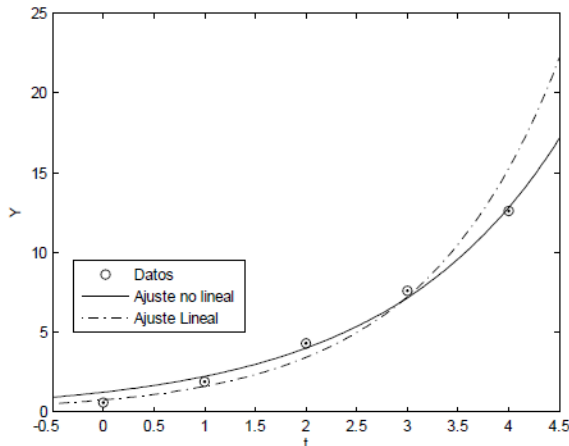
En la figura se aprecia que aplicando el ajuste no lineal se obtiene un mejor resultado que aplicando del ajuste lineal mediante una transformación logarítmica.

**Ejemplo 3.2** Se está buscando ajustar la función

$$g(t) = \frac{a}{1 + be^{tc}}$$

con los datos.

**Figura 1:** Gráfica de resultados obtenidos aplicando ajuste lineal y no lineal.



**Tabla 3.2.1** Datos.

$t_i$	$y_i$	$t_i$	$y_i$
1	5,308	7	31,443
2	7,240	8	38,558
3	9,638	9	50,156
4	12,866	10	62,948
5	17,069	11	75,995
6	23,192	12	91,972

Consideremos  $x_1 = a$ ,  $x_2 = b$ ,  $x_3 = c$ ; y definimos  $F = (F_i) \mathbb{R}^3 \rightarrow \mathbb{R}^{12}$  cuyas componentes están dados por

$$F_i(x) := \frac{x_1}{1 + x_2 \exp(t_i x_3)} - y_i, \quad i = 1, 2, \dots, 12.$$

Luego hay que solucionar el problema

$$\text{mín } f(x) = \|F(x)\|_2.$$

En este caso se tiene

$$F'(x) = \begin{pmatrix} \nabla_x F_1(x)^T \\ \nabla_x F_2(x)^T \\ \vdots \\ \nabla_x F_{12}(x)^T \end{pmatrix} \quad \text{donde } \nabla_x F_i(x) = \begin{pmatrix} 1 \\ \frac{x_1 \exp(t_i x_3)}{(1 + x_2 \exp(t_i x_3))^2} \\ -\frac{x_1 x_2 t_i \exp(t_i x_3)}{(1 + x_2 \exp(t_i x_3))^2} \end{pmatrix}, \quad i = 1, 2, \dots, 12.$$

Para el punto inicial  $x_0 := (200, 30, -0,4)^T$  y  $\delta = 0;0001$  en la prueba de Armijo y con  $\varepsilon = 10^{-12}$  en el

criterio para finalizar ( $f(x^k) - f_k(p_k) < \varepsilon$ ), se obtiene la siguiente corrida

**Tabla 3.2.2** Resultados del ejemplo 1.

$k$	$x_1^k$	$x_2^k$	$x_3^k$
0	200,00000000000000	30,00000000000000	-0,40000000000000
1	141,80746504198396	31,75257702369791	-0,34448829863712
2	171,20291006881448	40,80614279114224	-0,31029874032756
3	195,25942267327866	48,49540277681253	-0,31299183579093
4	196,16144824060422	49,08592600632490	-0,31358302855479
5	196,18593258549575	49,09159233284001	-0,31356989262609
6	196,18625897259517	49,09163901898217	-0,31356973125702

**Tabla 3.2.3** Resultados del ejemplo 1.

$k$	$f(x^k)$	$f_{x^k}(p_k)$	$f(x^k) - f_{x^k}(p_k)$
0	153,578482657057460	2,738461020895447	150,840021636162020
1	18,100388740764007	1,804975639384255	16,295413101379751
2	7,200408864389492	1,612677362162174	5,587731502227317
3	1,628737850586659	1,608496670492882	0,020241180093777
4	1,608511524594092	1,608501608095405	0,000009916498687
5	1,608501599487903	1,608501599404414	0,00000000083489
6	1,608501599403693	1,608501599403686	0,000000000000006

Se obtiene una aproximación de la solución óptima  $x^* \approx x^6$  y el valor óptimo

$$f(x^*) \approx \|F(x^6)\|_2 = 1,60850159940369.$$

Así se obtiene

$$a := x_1^6 = 196,18625897259517,$$

$$b := x_2^6 = 49,09163901898217,$$

$$c := x_3^6 = -0,31356973125702,$$

con el coeficiente de determinación múltiple  $R^2 = 0;99988411440177$ .

Análisis de varianza para probar la significación de la regresión:

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	$F_0$
Regresión	22323,54672768805200	3	7441,18224256268420	31636,73319969
Error o residual	2,58727739528424	11	0,23520703593493	
Total	22326,13400508333500	11		

Donde  $n=12$ ,  $p=1$ ;  $k=3$ . Para  $\alpha=0,05$ , como  $F_0=31636;73319969 > f_{\alpha,k,n-p}=3,59$  se rechaza la hipótesis nula.

**Ejemplo 3.3** Se está buscando ajustar la función  $g(t) = a + b e^{ct} + d e^{ft}$  con los datos

**Tabla 3.3.1** Datos.

$i$	1	2	3	4	5	6	7	8	9
$t_i$	0,0	0,5	1,0	1,5	2,0	3,0	5,0	8,5	10,0
$y_i$	3,85	2,95	2,63	2,33	2,24	2,05	1,82	1,80	1,75

Consideramos  $x_1=a, x_2=b, x_3=c, x_4=d, x_5=f$  y definimos  $F = (F_i) : \mathbb{R}^5 \rightarrow \mathbb{R}^{12}$  cuyas componentes están dados por

$$F_i(x) = x_1 + x_2 \exp(x_3 t_i) - y_i, i = 1, 2, \dots, 9$$

Luego hay que solucionar el problema

$$\min f(x) = \|F(x)\|_2.$$

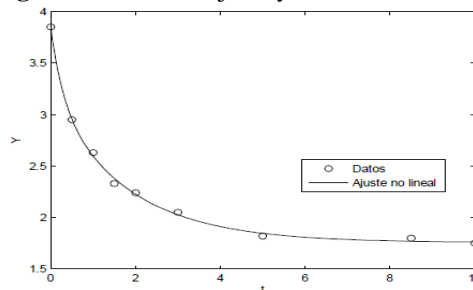
Se tiene

$$F'(x) = \begin{pmatrix} \nabla_x F_1(x)^T \\ \nabla_x F_2(x)^T \\ \vdots \\ \nabla_x F_9(x)^T \end{pmatrix}, \quad \nabla_x F_i(x) = \begin{pmatrix} 1 \\ \exp(x_3 t_i) \\ x_2 t_i \exp(x_3 t_i) \\ \exp(x_3 t_i) \\ x_4 t_i \exp(x_3 t_i) \end{pmatrix}, \quad i = 1, 2, \dots, 9.$$

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	$F_0$
Regresión	49,22905251164523	4	12,30726312791131	17959,55853981
Error o residual	0,00616748835478	9	0,00068527648386	
Total	49,23522000000001	9		

Donde  $n=12, p=1, k=3$ . Para  $\alpha=0,05$ , como  $F_0=17959;55853981 > f_{\alpha,k,n-p} = 3,63$  se rechaza la hipótesis nula.

**Figura 2.** Curva de ajuste  $y = a + b e^{c t} + d e^{f t}$ .



## 5. Conclusiones

En el algoritmo propuesto para solucionar el problema de aproximación no lineal, el problema es tratado como un caso especial de minimización no lineal sin restricciones, se aprovecha la estructura especial de la función objetivo. Para solucionar este problema en cada paso se tiene que determinar una dirección descendente y una longitud de paso para definir el siguiente punto iterativo.

Para obtener una dirección descendente se resuelve un problema de aproximación lineal, el cual es solucionado aplicando una factorización QR mediante las desviaciones de Householder. La longitud de paso se determina aplicando el algoritmo de búsqueda de Armijo.

Para la factorización QR se ha tomado en cuenta un criterio para evitar desbordamiento por exceso y por defecto (overflow y underflow).

Los algoritmos requieren que se introduzcan los jacobianos de cierta función.

Para el punto inicial  $x_0=(1,75, 1,20,-0,5, 0,8,-2,0)^T$  y  $\delta=0;0001$  en la prueba de Armijo y con  $\varepsilon = 10^{-12}$  en el criterio para finalizar ( $f(x^k)-f_x^k(p_k) < \varepsilon$ ), se obtiene una aproximación de la solución óptima  $x^* \approx x^8$  y el valor óptimo

$$f(x^*) \approx \|F(x^8)\|_2 = 0,07709708522938$$

Luego se considera

$$a := x_1^8 = 1,75773942237935,$$

$$b := x_2^8 = 1,42101552374600,$$

$$c := x_3^8 = -0,55525001434537,$$

$$d := x_4^8 = 0,67066466945013,$$

$$f := x_5^8 = -3,38357403802654,$$

con el coeficiente de determinación múltiple  $R^2=0;99987473421760$ .

Análisis de varianza para probar la significación de la regresión:

El ajuste no lineal es recomendable en caso de conocerse el modelo y en otro caso el investigador podría proponer un nuevo modelo. En este sentido, el ajuste no lineal es más flexible que el ajuste lineal.

Se han implementado los algoritmos propuestos para solucionar problemas de aproximación no lineal aplicando el método de Gauss-Newton y mediante los ejemplos presentados se puede apreciar su eficiencia y que su manejo es posible.

La convergencia del método dependerá del cumplimiento de las hipótesis por parte de la función objetivo y del punto inicial. En la práctica no se realiza la verificación de las hipótesis de convergencia, no son fáciles de verificar, sino simplemente se da un punto inicial y se inicializa la corrida. Un criterio para validar los resultados calculando es mediante el coeficiente de determinación múltiple  $R^2$ .

Algunos modelos de ajuste no lineal pueden ser transformados en un modelo lineal, sin embargo los

problemas de aproximación resultantes no son equivalentes y se tendrán soluciones distintas. En problemas de ajuste se recomienda solucionar los problemas sin transformarlos.

## **6. Referencias bibliográficas**

- Bersekas, D. P (1995). *Nonlinear Programming*, Athena Scientific, Belmont, USA.
- Burden, R. L. (1998). *Análisis numérico*. International Thomson Editores, International Thomson, Sexta Edición.
- Gerald, C. F.; Wheatley, P.O. (2000). *Análisis numérico con aplicaciones*. Pearson Educación. Sexta edición, México.
- Dennis, J. E, Jr ; Schnabel, R. B. (1996). *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, SIAM.
- Drubulle, A. A. (2000). Householder transformations revised, *Siam J. Matrix Anal. Appl.*, Vol 22, No. 1, pp. 33-40, April.
- Houch, P. D.; Vavasis, S. A. (1997). Complete orthogonal decomposition for weighted Leastsquares., *Siam J. Matrix Anal. Appl.*, Vol 18, No. 2, pp. 369-392, April.
- Montgomery, D. C.; Runger, G. C. (2006). *Probabilidad y Estadística Aplicadas a la Ingeniería*, Edotorial LIMUSA.
- Nocedal, J; Wright, S. J. (1999). *Numerical Optimization*, Springer-Verlag.
- Stoer, J; Bulirsch, R. (1993). *Introduction to Numerical Analysis*, Springer-Verlag.