# TEORÍA Y PRÁCTICA EN OPTIMIZACIÓN NUMÉRICA

*José Mario Martínez*

Novembro RP 51/92



## Relatório de Pesquisa

Instituto de Matemática
Estatística e Ciência da Computação

UNIVERSIDADE ESTADUAL DE CAMPINAS
Campinas - São Paulo - Brasil

**RESUMEN** - Analizamos las relaciones entre Práctica y Teoría en el área de Optimización Numérica. Consideramos problemas teóricos en grados crecientes de desafío y damos ejemplos de nuestra propia experiencia. Conjeturamos los campos en los cuales se desenvolverán teorías explicativas en el siglo XXI.

**ABSTRACT** - We analyze the relation between Practice and Theory in Numerical Optimization. We consider different degrees of challenge in current theoretical problems, and we give examples from our own experience. We conjecture which fields will provide explicative theories in the 21-th Century.

IMECC - UNICAMP
Universidade Estadual de Campinas
CP 6065
13081 Campinas SP
Brasil



**Novembro - 1992**

# Teoría y Práctica en Optimización Numérica

José Mario Martínez


## Resumen

Analizamos las relaciones entre Práctica y Teoría en el área de Optimización Numérica. Consideramos problemas teóricos en grados crecientes de desafío y damos ejemplos de nuestra propia experiencia. Conjeturamos los campos en los cuales se desenvolverán teorías explicativas en el siglo XXI.

## Abstract

We analyze the relation between Practice and Theory in Numerical Optimization. We consider different degrees of challenge in current theoretical problems, and we give examples from our own experience. We conjecture which fields will provide explicative theories in the 21-th Century.


---

Departamento de Matemática Aplicada, IMECC-UNICAMP, CP 6065, 13081 Campinas SP, Brasil.
Trabajo financiado por FAPESP, Projecto Temático 90/3724/6.

# 1. La Práctica tal como se presenta

La Optimización Numérica, como la entendemos en este artículo, se ocupa del problema de minimizar funciones continuas en dominios definidos por ecuaciones e inecuaciones:

$$\left.\begin{array}{ll} \text{Minimizar} & f(x) \\ \text{sujeita a} & h_1(x) = 0 \\ & h_2(x) \geq 0 \end{array}\right\} \tag{1.1}$$

donde $f : I\!R^n \to I\!R$, $h_1 : I\!R^n \to I\!R^m$, $h_2 : I\!R^n \to I\!R^p$.
Varios casos particulares de (1.1) merecen atención especial. Cuando no hay restricciones, el dominio donde se desea minimizar $f$ es $I\!R^n$, y hablamos, en ese caso, de Minimización Irrestricta. Cuando $f, h_1$ y $h_2$ son lineales, se trata del problema de Programación Lineal. Si $f$ es cuadrática, $h_1$ y $h_2$ lineales, (1.1) es el problema de Programación Cuadrática, etc.
Un marco muy importante para el análisis y la resolución de (1.1) es el problema de resolver un sistema no lineal, o sea, encontrar $x \in I\!R^n$ tal que

$$F(x) = 0 \tag{1.2}$$

donde $F : I\!R^n \to I\!R^n$. Efectivamente, por un lado las condiciones de optimalidad de (1.1) son de tipo (1.2) y, por otro lado, (1.2) es en sí mismo un problema de optimización, a través de su transformación en

$$\text{Minimizar} \quad \|F(x)\|^2. \tag{1.3}$$

El objetivo último de la Optimización Numérica (O.N.) es el desarrollo de algoritmos computacionais eficientes para resolver problemas prácticos de tipo (1.1). No pretendemos dar una definición rigurosa de *Práctica* en el contexto de la O.N.. Antes bien, vamos a describir aquello que los optimizadores numéricos han pensado los últimos 35 años al usar ese término.

## 1.1 Problemas de juguete

La denominación "Problemas de juguete" (Toy problems) no tiene aquí contenido despectivo. Se trata del conjunto de problemas de descripción fácil, encontrables y a veces colecionados en la literatura, y raras veces vinculados a aplicaciones reales. El problema más famoso de ese tipo es la minimización de la función de Rosenbrock, $f(x_1, x_2) = 100\ (x_2 - x_1^2)^2 + (1 - x_1)^2$. Algunos problemas de juguete fueron inventados con dificuldades específicas pero, frecuentemente, el objetivo inicial de la invención fue olvidado. El mérito principal de los problemas de juguete es

ofrecer un conjunto de tests facilmente reproducibles que sirven para dar una idea preliminar sobre el funcionamiento de algoritmos nuevos.
La colección más popular de problemas de juguete en Minimización sin Restricciones es la publicada por Moré, Garbow y Hillstrom [1981]. Se trata de 35 problemas, 15 de los cuales de dimensión variable, y los otros 20 con dimensiones de hasta 31 variables. Unos pocos de estos problemas se originaron en aplicaciones, pero la gran mayoría son artificiales.
Para Minimización con Restricciones, la colección más difundida de problemas de juguete es la publicada por Hock y Schittkowski [1981].

## 1.2 Colecciones de Problemas Reales

Los esfuerzos por coleccionar problemas de optimización provenientes de aplicaciones reales son recientes. Ver Moré [1989], Averick, Carter y Moré [1991], Conn, Gould y Toint [1992].
La colección de Moré [1989] incluye:

1) Problemas de autovalores con peso positivo, con aplicación en niveles de energía, estados cuánticos de átomos y moléculas.

2) Determinación de las condiciones de equilibrio de columnas de destilación.

3) Fenómenos de difusión no lineal en combustión y semiconductores.

4) Ecuación Integral de Chandrasekhar.

5) Problemas de lubricación Elastohidrodinámica.

6) Flujo Renal.

7) Problemas de Obstaculo.

8) Modelo de reactor químico tubular.

9) Modelo de estabilidad de aeroplanos.

10) Problemas de semiconductores.

En Averick, Carter y Moré [1991] aparecen los siguientes problemas:

1) Inyección de flúido por un lado de un canal vertical largo.

2) Flúidos en remolino.

3) Determinación experimental electrolítica del momento dipolar en el corazón humano.

4) Combustión de Propano.

5) Isomerización de $\alpha$ - pineno.

6) Detección no destructiva de falta de uniformidad en revestimientos.

7) Problemas de ajustes exponenciales de datos.

8) Análisis de la resistencia de un termostato como función de la temperatura.

9) Análisis de una reacción enzimática.

10) Torsión elastoplástica.

11) Distribución de presiones en una película de lubricante entre dos cilindros circulares.

12) Diseño óptimo con materiales compuestos.

13) Superconductores no homogéneos.

Las colecciones de problemas reales son útiles porque, probablemente, representam más adecuadamente que los problemas de juguete, las aplicaciones que se encuentran en Física, Química, Ingeniería, etc. Sin embargo, como veremos más adelante, desde que no sabemos exatamente qué significa "representar", también estos testes deben ser considerados apenas como "parte" de la Práctica cotidiana en Optimización Numérica.
Todavía hay dificultades de comunicación y reproducción de este conjunto de problemas. Las funciones que los definen no son tan simples como en el caso de problemas de juguete, y, por lo tanto, su difusión a través de la literatura está sujeta a bastantes errores. Una manera de eliminar, o disminuír, estos errores, es difundir los problemas mediante la transmisión directa de los programas, por correo electrónico. Esa tarea está siendo ejecutada con mucho éxito para problemas de Programación Lineal (Gay [1985]). En este caso, la definición del problema envuelve apenas datos numéricos (no funciones) y el formato MPS de IBM es aceptado en el mundo entero. Para problemas no lineales la situación no es tan simple. En este caso, la definición del problema envuelve funciones y, por lo tanto, sería necesario una convención que

unificase lenguajes y características de las subrutinas, así como parámetros, estructuras de datos, y áreas en common que las definen.

## 1.3 Problemas Aleatorios

Generando aleatoriamente los datos, pueden ser creados infinitos problemas diferentes, algunas de cuyas características pueden ser fijadas a priori. Esta técnica ha sido usada frecuentemente para generar problemas de Programación Cuadrática:

$$\left.\begin{array}{ll} \text{Minimizar} & \frac{1}{2}x^T G x + c^T x \\ \text{sujeita a} & Ax = b \\ & \quad x \geq 0 \end{array}\right\} \tag{1.4}$$

Se trata de producir problemas de tipo (1.4) donde la solución y otras particularidades sean conocidas. Las etapas de esa tarea son las siguientes:

1) Generar la matriz $A$, con estructura predeterminada, densidad de los elementos no nulos, etc.

2) Generar una solución $x_*$, decidiendo de antemano componentes iguales a cero (grado de degeneración primal).

3) Calcular $b$, usando $A$ y $x_*$.

4) Generar $G$, predeterminando autovalores, estructura y densidad.

5) Decidir número de multiplicadores de Lagrange nulos (grado de degeneración dual), y generar los restantes.

6) Usando $G$ y los multiplicadores de Lagrange, calcular c.

La ventaja de este sistema es que se pueden generar muchos problemas con una propiedad dada y, consecuentemente, testear la eficiencia de un algoritmo ante Hessianos singulares, grados de singularidad, de degeneración, etc.
Sin embargo, muchas veces el desempeño de un algoritmo en problemas reales resulta totalmente diferente de lo previsto por los experimentos aleatorios. Esto se debe a que la estructura de los problemas aleatorios es "exactamente" la inducida por el experimentador, mientras que en los problemas reales hay muchas más características estructurales no percibidas. No sabemos, inclusive, si poseemos ya herramientas matemáticas útiles para describir tales *estructuras ocultas* pero, lo que es cierto es que ellas no aparecen en experimentos aleatorios, cuya esencia es, precisamente, destruír

todas las estructuras no deseadas.

## 1.4 Problemas de Consultoría

Muchos problemas reales admiten ser analizados mediante Modelos Matemáticos. Éstos son problemas matemáticos, cuya resolución es interpretable como solución, por lo menos aproximada, del problema original. En nuestro caso, entenderemos que el problema matemático, o modelo, es de Optimización. En un primer momento, el consultor optimizador recibe un modelo matemático para ser resuelto, o participa de su formulación con el "consultante", digamos, un físico. La formulación del modelo, o sea, el tipo de problema matemático, es también influenciada por las técnicas de resolución disponibles. Una vez aceptado el modelo por ambas partes, el optimizador procede a su resolución usando la batería de algoritmos que considera adecuada. En general, tiene éxito en esta tentativa. Sin embargo, en la inmensa mayoría de los casos, el físico rechaza esta primera solución por considerarla imposible en relación a su problema real. Normalmente, el físico mostrará al consultor que tal variable no puede crecer tan rápido, tal otra no puede ser negativa, y que determinada relación entre algumas variables no se cumple. El optimizador reconocerá inmediatamente que lo ocurre es que las restricciones que el físico está acusando como violadas no estaban en el modelo original, razón por la cual, justamente, no fueron respetadas. Proceden, en consecuencia, a modificar el modelo, teniendo en cuenta las nuevas restricciones. Esta modificación puede exigir la aplicación de técnicas diferentes de las utilizadas originalmente.

El processo de ida y vuelta esbozado arriba puede repetirse muchas veces. Frecuentemente "no converge", a veces por razones no estrictamente científicas. Sin embargo, sea cual fuere su conclusión, deja un residuo de experiencia práctica que, contrariamente a la reseñada en las subsecciones anteriores, no es documentada ni coleccionada de forma alguma. Paradójicamente, es bastante difundida la opinión de que ésta es la experiencia que realmente importa, seguramente porque es el único tipo vinculado a flujos económicos.

A veces, el consultor optimizador acumula un enorme bagaje de experiencia, mas apenas en una única clase de problemas. Por ejemplo, durante muchos años ha resuelto problemas de Ingeniería Química, probablemente porque trabaja en un Departamento de esa área en una universidad o industria. De hecho, en esos casos, el consultor es más identificado por su área de aplicación que por su especialidad dentro de la matemática. La experiencia de este tipo de consultor también es muy importante y, en general, permanece igualmente no documentada.

Los consutores son depositarios, pues, de rica e influyente experiencia para el desarrollo de algoritmos. La subjetividad y falta de comunicación confiable son los obstáculos obvios para la generalización de su influencia entre los inventores de algoritmos.

## 2. El Desafío Teórico

Nos ocupamos aquí de *Teoría Matemática sobre Problemas y Algoritmos de Optimización*. Es casi imposible que una teoría matemática hable de un algoritmo tal como es en la práctica. De hecho, en teoría trabajamos como si los algoritmos generasen sucesiones infinitas de puntos en $\mathbb{R}^n$, cuando en la realidad generamos apenas una sucesión bastante corta de puntos en un conjunto finito (el de las n-uplas representables en la computadora ) mucho menor que $\mathbb{R}^n$. Por lo tanto, un teorema rara vez se aplica *directamente* a la predicción de un comportamiento numérico. Se aplica, sí, a través de un trabajo de *Interpretación*.

Las teorías matemáticas sobre algoritmos y problemas son corretas dentro del universo matemático, independientemente de los experimentos. Las teorías son útiles porque, cuando correctamente *interpretadas* sirven para predecir el desempeño de algoritmos o la posibilidad de resolver ciertos problemas y, por lo tanto, orientar el desarollo de nuevas técnicas de manera mucho más económica que a través de ensayos y errores. La relevancia, la motivación y el objetivo de teoría matemática en Optimización está vinculada a su interpretabilidad en términos de desempeño práctico de métodos. Curiosidades o ejercicios de Análisis Matemático pueden, eventualmente, ser interesantes, pero no en este contexto.

Podemos ordenar las posibles teorías matemáticas em Optimización Numérica de acuerdo con una jerarquía de creciente desafío e interés, medido por el posible impacto práctico en el desarollo de algoritmos. Para simplificar, mencionaremos apenas tres escalones de relevancia:

1) Teoría que, de acuerdo con *interpretación* ya consagrada, explica el comportamiento práctico de algoritmos.

2) Teoría destinada a explicar comportamientos para los cuales no hay, por el momento, explicación teórica disponible.

3) Teoría destinada a explicar comportamientos *discrepantes* con otras teorías existentes, de acuerdo con las interpretaciones usuales.

Podemos observar que lo que va siendo crecientemente cuestionado en nuestra jerarquía es precisamente la *interpretación*. En el primer grupo, la interpretación corriente no es cuestionada, y la teoría se desarrolla dentro de paradigma consagrado, sin grandes conflictos. En el segundo grupo, las interpretaciones son insuficientes, pero todavía no están cuestionadas. El tercer grupo es el más interessante porque, aparentemente, nos lleva a revisar el significado (interpretación) habitual dado a una teoría.

En las secciones siguientes, daremos ejemplos de los tres tipos de cuestiones teóricas, tomados de nuestra propria experiencia.

## 3. Huyendo de Puntos de Ensilladura

Recientemente, Dennis, Echebest, Guardarucci, Martínez, Scolnik y Vaccino [1991] introdujeron un algoritmo para minimización sin restricciones basado en búsquedas curvilíneas en el camino Levenberg-Marquardt. El algoritmo básico es el siguiente.

**Algoritmo 3.1** Sea $\Delta$ un numero positivo (grande), $\alpha \in (0,1/2)$, $\beta \in (0,1)$. Dado $x_k \in \mathbb{R}^n$ tal que $\nabla f(x_k) \neq 0$, $H_k$ simétrica com autovalores $\lambda_1(H_k) \leq \ldots \leq \lambda_n(H_k)$, definimos

$$\psi_k(x) = \frac{1}{2}(x - x_k)^T H_k(x - x_k) + \nabla f(x_k)^T(x - x_k), \tag{3.1}$$

$$\begin{aligned} \Gamma(x_k, H_k) = \{ \; & x \in \mathbb{R}^n \text{ tales que } x \text{ minimiza} \\ & \psi_k(x) \text{ en el conjunto } \|x - x_k\| \leq \delta, \\ & \text{para algún } \delta > 0\} \end{aligned} \tag{3.2}$$

Los pasos para calcular $x_{k+1}$ son:

*Paso 1.* Se $\lambda_1(H_k) \leq 0$, definir $\Delta_k \leftarrow \overline{\Delta}_k \leftarrow \Delta$. Si no, definir $s_k^N = -H_k^{-1}\nabla f(x_k)$, $\Delta_k \leftarrow \overline{\Delta}_k \leftarrow \min\{\Delta, \frac{\|S_k^N\|}{\beta}\}$.

*Paso 2.* Elegir $\overline{x} \in \Gamma(x_k, H_k)$ tal que

$$\beta^2 \Delta_k \leq \|\overline{x} - x_k\| \leq \Delta_k. \tag{3.3}$$

*Paso 3.* Si

$$f(\overline{x}) \leq f(x_k) + \alpha\psi_k(\overline{x}), \tag{3.4}$$

definir $x_{k+1} = \overline{x}$.

*Paso 4.* $\Delta_k \leftarrow \Delta_k/2$.

De hecho, el algoritmo publicado en Dennis et al [1991] es ligeramente diferente de éste. Allí, en vez de la condición (3.4), tenemos la condición "de Armijo"

$$f(\overline{x}) \leq f(x_k) + \alpha \ \nabla f(x_k)^T(x - x_k). \tag{3.5}$$

En la implementación computacional, hacemos $\alpha = 10^{-4}$ y elegimos $\overline{x}$ como un minimizador aproximado de $f$ en $\Gamma(x_k, H_k)$, lo que es, en general, bastante más exigente que (3.4) o (3.5). Como consecuencia, el desempeño práctico de la implementación basada en (3.4) es idéntico al basado en (3.5).

En Dennis et al [1991] fue demostrado el siguiente teorema.

**Teorema 3.2** *Si existe $M > 0$ tal que $\|H_k\| \leq M$ para todo $k$, entonces todo punto de acumulación $x_*$ de una sucesión $\{x_k\}$ generada por el Algoritmo 3.1 es estacionario de primer ordem, o sea, satisface*

$$\nabla f(x_*) = 0. \tag{3.6}$$

Resultados como el enunciado en el Teorema 3.2 son típicos en la teoría de minimización sin restricciones. Es generalmente admitido que algoritmos que pueden generar puntos de acumulación no estacionarios no merecem siquiera ser sometidos a experimentación.

En la implementación del Algoritmo 3.1 elegimos $H_k = \nabla^2 f(x_k)$ siempre que $k$ es múltiplo de un entero fijo $q$. En los otros casos, usamos una fórmula Quasi-Newton para construír $H_k$.

Como es bien sabido, puntos $x_*$ que satisfacen (3.6) pueden no ser minimizadores locales de $f$. De hecho, si el Hessiano $\nabla^2 f(x_*)$ tiene algún autovalor negativo, $x_*$ no será minimizador local. Algunas funciones test de juguete tienen puntos estacionarios donde el Hessiano tiene autovalores positivos y negativos. Algoritmos que satisfacen *apenas* el Teorema 3.2 tiendem, según el caso, a converger a esos puntos estacionarios (llamados "de ensilladura") o, más frecuentemente, a pasar muchas iteraciones en los alrededores de esos puntos, antes de conseguir apartarse definitivamente de su influencia. Sin embargo, en nuestra implementación del Algoritmo 3.1, nada de eso fue observado. El algoritmo no mostró ninguna atracción por puntos de ensilladura que no fueram verdaderos minimizadores, lo que nos llevó a conjeturar que algún resultado teórico más fuerte debía ser válido. Así, en Martínez [1992a], demostramos el siguiente teorema.

**Teorema 3.3** *Supongamos que* $f \in C^2(\mathbb{R}^n)$, *y que la sucessión infinita* $\{x_k\}$ *es obtenida usando el Algoritmo 3.1. Supongamos que* $K_1 \subset \mathbb{N}$ *es un conjunto infinito de índices tal que* $\lim_{k \in K_1} x_k = x_*$ *y* $\lim_{k \to K_1} H_k = \nabla^2 f(x_*)$. *Entonces* $\nabla f(x_*) = 0$ *y* $\nabla^2 f(x_*)$ *es positiva semidefinida.*

Esencialmente, el Teorema 3.3 dice que subsucesiones de la sucesión generada por nuestro algoritmo no pueden converger a puntos estacionarios donde el Hessiano tiene algún autovalor negativo. La interpretación de este resultado teórico es que dificilmente la solución arrojada por el método en un caso práctico será un punto de ensilladura, y que la sucesión pasará rapidamente por la proximidad de esos puntos, sin sentirse tentada por sus cantos de sirena. Este comportamiento es, precisamente, el observado en la práctica. La interpretación natural de la teoría fue así confirmada, luego, estamos en presencia de resultados teóricos de tipo 1, de acuerdo con la clasificación de la Sección 2.

# 4. El Método de Actualización de Columna

Desde la década del 60 los métodos Quasi Newton se cuentam entre los más eficientes para resolver Sistemas No Lineales y otros problemas de Optimización.
Consideremos el problema de resolver

$$F(x) = 0, \tag{4.1}$$

donde $F : \mathbb{R}^n \to \mathbb{R}^n$ es diferenciable. La iteración típica de un método Quasi Newton es

$$x_{k+1} = x_k - B_k^{-1} F(x_k). \tag{4.2}$$

Definiendo $s_k = x_{k+1} - x_k$, $y_k = F(x_{k+1}) - F(x_k)$, los métodos Quasi Newton que satisfacen

$$B_{k+1} s_k = y_k \tag{4.3}$$

para todo $k \in \mathbb{N}$ se llaman *secantes.* Las teorías de convergencia para métodos secantes fueron ganando amplitud en los últimos 20 años. Ver Martínez [1990, 1992d] y los artículos previos de Broyden, Dennis y Moré [1973], Dennis y Moré [1977], Dennis y Walker [1991], etc.
El método de Broyden (Broyden [1965]) es uno de los métodos secantes más utilizados. En ese método, la generación de $B_{k+1}$ a partir de $B_k$ corresponde a la fórmula

$$B_{k+1} = B_k + \frac{(y_k - B_k s_k) s_k^T}{s_k^T s_k}. \tag{4.4}$$

Por (4.4) y por la fórmula de Sherman y Morrison (Golub y Van Loan [1989]) también tenemos que,

$$B_{k+1}^{-1} = (I + u_k s_k^T) \dots (I + u_0 s_0^T) B_0^{-1} \quad , \tag{4.5}$$

para todo $k \in \mathbb{N}$, donde $z_k = s_k$ y

$$u_k = \frac{(s_k - B_k^{-1} y_k)}{s_k^T B_k^{-1} y_k} \quad . \tag{4.6}$$

De (4.5) resulta que el método de Broyden puede ser implementado usando $O(n^2)$ operaciones por iteración, cuando $n$ es pequeño, y que una implementación para problemas grandes también es posible (Gomes-Ruggiero, Martínez y Moretti [1992]).

Con la intención de desarollar un método secante con propiedades similares a Broyden, pero con una fórmula de actualización aún más barata que (4.5), introdujimos el Método de Actualización de una Colunna por Iteración (CUM) (Martínez [1984]). En CUM, en vez de (4.4), tenemos

$$B_{k+1} = B_k + \frac{(y_k - B_k s_k) e_{j_k}^T}{e_{j_k}^T s_k} \tag{4.7}$$

donde $e_{j_k}$ es un vector de la base canónica de $\mathbb{R}^n$ para el cual $|e_{j_k}^T s_k| = \|s_k\|_\infty$. Consecuentemente, las fórmulas (4.5) y (4.6) continúan válidas para CUM, con $z_k = e_{j_k}$. El costo en tiempo y memoria de CUM es menor que el de Broyden, aunque del mismo orden. Hasta 1991, los principales resultados de convergencia para Broyden y CUM eram los siguientes.

**Teorema 4.1** *Supongamos que $F(x_*) = 0, F'(x_*)$ no singular, y que existem $L, p > 0$ tales que*

$$\|F'(x) - F'(x_*)\| \leq L \|x - x_*\|^p \tag{4.8}$$

*para todo $x$ en un entorno convexo de $x_*$. Entonces existen $\varepsilon, \delta > 0$ tales que, si $\|x_0 - x_*\| \leq \varepsilon$ y $\|B_0 - J(x_k)\| \leq \delta$, la sucesión $\{x_k\}$ generada por Broyden está bien definida, converge a $x_*$ y satisface*

$$\lim_{k \to \infty} \frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|} = 0 \quad . \tag{4.9}$$

*Demostración.* Ver Dennis y Schnabel [1983]. La propiedad (4.9) se denomina *Convergencia Q-superlineal.* □

**Teorema 4.2** *Supongamos las mismas hipótesis sobre $F$ que en el Teorema 4.1. Consideremos el método CUM con la siguiente modificación: si $k+1$ es múltiplo de un entero fijo $m$, elegimos $B_{k+1} = F'(x_k)$. Entonces, existe $\varepsilon > 0$ tal que, si $\|x_0 - x_*\| \leq \varepsilon$, la sucesión $\{x_k\}$ está bien definida, converge a $x_*$, y satisface (4.9).*

*Demostración.* Ver Martínez [1984]. Una extensión para espacios de Hilbert se encuentra en Gomes-Ruggiero y Martínez [1992]. Un teorema análogo para la "versión inversa" de CUM se halla em Martínez y Zambaldi [1992]. □

Existe uma gran distancia entre el resultado del Teorema 4.1 y el del Teorema 4.2. El Teorema 4.1 nos dice que la convergencia local superlineal para Broyden se obtiene, para un punto inicial suficientemente cercano de $x_*$, exigiendo apenas que $B_0$ esté bastante próximo de $F'(x_*)$,o, en la práctica, de $F'(x_k)$. El Teorema 4.2 dice que el mismo resultado se obtiene para CUM pero exigiendo mucho más: que *todas* las aproximaciones $B_k$, cuando $k$ es múltiplo de $m$, sean Jacobianos. La debilidad del Teorema 4.2 se ve confirmada por el hecho de que el mismo resultado vale para el "Método de Newton estacionario con recomienzos periódicos", en el cual $B_{k+1} = B_k$ si $k$ no es múltiplo de $m$ y $B_{k+1} = F'(x_{k+1})$ si lo es.

No sabemos todavia si el método CUM *puro*, es decir, sin los recomienzos periódicos, tiene ni siquiera convergencia local.

A despecho de la debilidad de los resultados teóricos, realizamos a partir de 1984 cientos de experimentos numéricos usando Broyden y CUM. Los resultados son sorprendentes:

a) Jamás encontramos un caso donde uno de los dos métodos converge y el otro no.

b) La diferencia entre el número de iteraciones usado por uno y otro raramente excede un 10 por ciento del total.

c) En general, CUM usa menos tiempo que Broyden, debido al menor trabajo en Álgebra Lineal por iteración.

Nos encontramos, pues, en presencia de desempeños prácticos no explicados por la teoría existente. Varias veces retomamos la tarea de intentar cerrar el intervalo que separa aquí teoría y práctica. En principio, la tarea es tratar de encontrar resultados teóricos de convergencia más fuertes para CUM. Em 1992 conseguimos probar los siguientes teoremas.

**Teorema 4.3** *Bajo las hipótesis sobre $F$ del Teorema 4.1, consideramos el método CUM (sin recomienzos). Supongamos que la sucesión $\{x_k\}$ está bien definida, converge a $x_*$, y existe $r > 0$ tal que*

$$\|x_{k+1} - x_*\| \leq r \; \|x_k - x_*\| \tag{4.10}$$

*para todo $k \in I\!N$. Entonces*

$$\lim_{k\to\infty} \frac{\|x_{k+2n} - x_*\|}{\|x_k - x_*\|} = 0 \tag{4.11}$$

*y*

$$\lim_{k\to\infty} \|x_k - x_*\|^{1/k} = 0 \ . \tag{4.12}$$

*Demostración.* Ver Martínez [1992c]. La propiedad (4.11) se denomina *Convergencia 2n-Q-superlineal*, y (4.12) es *convergencia R-superlineal.* □

**Teorema 4.4** *Con las hipótesis habituales sobre $F$, consideramos el método CUM con la siguiente modificación: Dado $m \geq 2n$, si $k+1$ es múltiplo de $m$, la fórmula (4.7) puede no ser obedecida. Entonces existem $\varepsilon, \delta > 0$ tales que, si $\|x_0 - x_*\| \leq \varepsilon$ y $\|B_{k+1} - F'(x_*)\| \leq \delta$ para todo $k$ tal que $k+1$ es múltiplo de $m$, la sucesión $\{x_k\}$ está bien definida, converge para $x_*$ y satisface (4.12).*

*Demostración.* Ver Martínez [1992c]. □

**Teorema 4.5** *Con las mismas hipótesis sobre $F$, consideramos el método CUM, con $n = 2$. Entonces dado $r \in (0,1)$, existem $\varepsilon, \delta > 0$ tales que, si $\|x_0 - x_*\| \leq \varepsilon$ y $\|B_0 - J(x_*)\| \leq \delta$, la sucesión $\{x_k\}$ está bien definida, converge a $x_*$, y satisface (4.10), (4.11) y ( 4.12).*

*Demostración.* Ver Martínez [1992c]. □

Los Teoremas 4.3, 4.4 y 4.5 son avances claros en relación al Teorema 4.2. Por lo tanto, contribuyen, en parte, a explicar por qué los comportamientos de Broyden y CUM son tan parecidos. Sin embargo, el problema todavía está lejos de haber sido resuelto. En particular, la convergencia local de CUM *sin recomienzos* sólo fue demonstrada para $n = 2$, lo que significa poco si consideramos que Broyden y CUM tienen desempeños casi idénticos justamente en problemas grandes. La comparación entre (4.11) y (4.9) también es llamativa. Tomadas literalmente, estas propiedades parecen decir que CUM tarda 2n iteraciones para hacer algo que Broyden hace en una. Esto es muy diferente de lo observado en la práctica.
Por lo tanto, todavía existen muchos interrogantes teóricos para explicar la eficiencia de CUM. Creemos que estos interrogantes pueden ser formulados como problemas abiertos dentro de los paradigmas teóricos vigentes. Por lo tanto, nos parece que encaramos, en este caso, una teoría de tipo 2, de acuerdo con la clasificación de la

sección 2 de este artículo. En la próxima sección estudiaremos un caso del tercer tipo.

## 5. Discrepancia entre Teoría y Práctica

Grandes sistemas no lineales, problemas de minimización sin restricciones y de Programación No Lineal son muchas veces resueltos usando el Método de Newton Inexacto, a veces llamado Newton Truncado, que, para el caso de sistemas no lineales puede ser formulado de la siguiente manera: Dado $x_k \in \mathbb{R}^n$, la k-ésima aproximación para resolver (4.1), calcúlase $x_{k+1}$ definiendo $s_k$ tal que

$$\|F'(x_k)s_k + F(x_k)\| \leq \theta_k \|F(x_k)\| \tag{5.1}$$

y haciendo

$$x_{k+1} = x_k + s_k \quad , \tag{5.2}$$

para todo $k \in \mathbb{N}$, donde $\theta_k \in (0,1)$ . Como vemos, $s_k$ es una solución aproximada del sistema lineal newtoniano

$$F'(x_k)s = -F(x_k). \tag{5.3}$$

En la implementación, el incremento $s_k$ que satisface (5.1) se obtiene aplicando un método iterativo lineal al sistema (5.2), normalmente, métodos de tipo "subespacios de Krylov", categoría a la cual pertenecen el clásico método de gradientes conjugados (Hestenes y Stiefel [1952]) y el más reciente GMRES (Saad y Schultz [1986]). El siguiente teorema, de Dembo, Eisenstat y Steihaug describe la convergencia del método basado em (5.1) y (5.2).

**Teorema 5.1** *Supongamos que* $F(x_*) = 0$, $J(x_*)$ *no singular y continua en* $x_*$, *y* $\theta_k \leq \theta_{\max} < \theta < 1$ *para todo* $k \in \mathbb{N}$. *Existe* $\varepsilon > 0$ *tal que, si* $\|x_0 - x_*\| \leq \varepsilon$ *entonces la sucesión* $\{x_k\}$ *generada por (5.1) - (5.2) está bien definida, converge a* $x_*$, *y satisface*

$$\|x_{k+1} - x_*\|_* \leq \theta \|x_k - x_*\|_* \tag{5.4}$$

*para todo* $k \in \mathbb{N}$, *definiendo* $\|y\|_* = \|J(x_*)y\|$. *Si* $\lim_{k\to\infty} \theta_k = 0$, *la convergencia es Q-superlineal.*

*Demontración.* Ver Dembo, Eisenstat y Steihaug [1983]. □

Por el Teorema 5.1 y la condición (5.1), observamos que, para obtener convergencia Q-superlineal, necesitamos en principio que $\theta_k$ tienda a 0. O sea, el grado de precisión requerido en la solución iterativa de (5.3) debe ser cada vez mayor, y tender a infinito, cuando aumenta $k$. Por lo tanto, el número de pasos del método iterativo lineal que aplicamos a (5.3), también debe crecer mucho con $k$. Esto es un serio inconveniente, ya que la eficiencia del Método de Newton Inexacto está vinculada a que el costo de la iteración principal sea moderado.

Esto no sólo es verdadero para los sistemas lineales grandes provenientes de sistemas no lineales, sino también para sistemas lineales grandes en general, independientemente de su origen. Para que el número de iteraciones de un método iterativo aplicado a un sistema lineal sea pequeño, se suelen usar técnicas de Precondicionamiento. En el caso del sistema (5.3), *precondicionar* significa reemplazar (5.3) por el sistema equivalente

$$B_k^{-1}F'(x_k)s = -B_k^{-1}F(x_k), \tag{5.5}$$

donde $B_k^{-1}F'(x_k)$ debe ser una matriz bien condicionada, y $B_k^{-1}$ debe ser fácil de calcular. Existen muchas técnicas de precondicionamiento de sistemas lineales. La más popular es hacer $B_k$ igual a la matriz diagonal cuyos elementos son precisamente los de la diagonal del sistema original.
Recientemente, comezamos a preguntarmos si la particularidad de tener que resolver una sucesión de sistemas lineales *parecidos*, con la forma (5.3), sugeriría algún tipo de precondicionamiento a través del cual la exigencia $\theta_k \to 0$ en (5.1) no fuera necesaria. Así, llegamos a la formulación del siguiente algoritmo.

**Algoritmo 5.2** Sea $\theta_k \in (0,\theta)$ para todo $k \in I\!N, \theta \in (0,1)$ y $\lim\limits_{k\to\infty}\theta_k = 0$. Supongamos que $x_0 \in I\!R^n$, y que $B_0 \in I\!R^{n\times n}$, no singular, es el precondicionador inicial. Dado $x_k \in I\!R^n, B_k$ no singular, los pasos para obtener $x_{k+1}$ son los siguientes.

*Paso 1.* Calcular

$$s_k^Q = -B_k^{-1}F(x_k). \tag{5.6}$$

*Paso 2.* Si

$$\|F'(x_k)s_k^Q + F(x_k)\| \le \theta\|F(x_k)\| \tag{5.7}$$

definir

$$s_k = s_k^Q \tag{5.8}$$

si no, encontrar un incremento $s_k$ tal que (5.3) se satisface, usando algún método iterativo.

*Paso 3.* Definir $x_{k+1} = x_k + s_k$.

$\odot$

En el Algoritmo 5.2, $B_k$ es interpretado naturalmente como precondicionador de (5.3) debido a que, en cualquier método de subespacios de Krylov precondicionado aplicado a (5.3), $s_k^Q$ definido por (5.6) es necesariamente la primera iteración. Así, el Algoritmo 5.2 dice que, si la primera iteración del método iterativo lineal precondicionado satisface la condición débil (5.7), ella será aceptada como incremento. En caso contrario, pediremos la condición fuerte (5.1). El Teorema 5.3 enuncia el principal resultado de convergencia relativo al Algoritmo 5.2.

**Teorema 5.3** *Supongamos que $F : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$, $\Omega$ abierto y convexo, $F \in C^1(\Omega)$, $F'(x_*)$ no singular, y existem $L > 0$, $p \geq 1$ tales que $\|F'(x) - F'(x_*)\| \leq L\|x - x_*\|^p$. Supongamos que $\|B_k\|$ y $\|B_k^{-1}\|$ están acotados y que*

$$\lim_{k\rightarrow\infty} \frac{\|[B_k - F'(x_*)]s_k\|}{\|s_k\|} = 0 \ . \tag{5.9}$$

*Entonces existe $\varepsilon > 0$ tal que, si $\|x_0 - x_*\| \leq \varepsilon$, la sucesión $\{x_k\}$ generada por el Algoritmo 5.2 está bien definida y converge Q-superlinealmente a $x_*$. Más aún, existe $k_0 \in \mathbb{N}$ tal que $s_k = s_k^Q$ para todo $k \geq k_0$.*

*Demostración.* Ver Martínez [1992b]. $\square$

El Teorema 5.3 dice que, si usamos una sucesión de precondicionadores $B_k$ que satisfagan la condición (5.9), entonces obtenemos convergencia superlineal sin necesitar que $\theta_k$ tienda a cero, y tampoco necesitando que $B_0$ sea cercano a ningún Jacobiano verdadero. Una sucesión de precondicionadores que satisfacen (5.9) puede ser generada usando procedimientos definidos en Martínez [1992b]. Como caso particular, si $B_{k+1}^{-1}$ es construída a partir de $B_k^{-1}$ usando la fórmula de Broyden ((4.4) y (4.5)), podemos probar que $\|B_k\|$ y $\|B_k^{-1}\|$ son acotadas y que vale (5.9). Por lo tanto, el Teorema 5.3 resuelve de manera elegante la cuestión que nos planteamos después del Teorema 5.1. Por éste, parecía que la convergencia superlineal estaba asociada a $\theta_k \rightarrow 0$, o sea al crecimiento de trabajo en cada iteración del método de Newton Inexacto. En el Teorema 5.3 vemos que, si usamos precondicionadores con una cierta característica, obtenemos convergencia superlineal haciendo apenas una iteración a partir de $k_0$. En otras palavras, sin necesidad de que el primer condicionador $B_0$ sea bueno, los precondicionadores $B_k$ van ganando información a lo largo del proceso, hasta que, para $k$ suficientemente grande, el paso $-B_k^{-1}F(x_k)$ es capaz, por sí solo, de garantizar la convergencia superlineal.

La *Interpretación* del Teorema 5.3 en términos prácticos nos pareció, al principio, muy clara: Si usásemos la fórmula de Broyden, por ejemplo, como precondicionador, definiendo $\theta = \frac{1}{2}$, $\theta_k = \frac{1}{k+1}$ para todo $k$, con $B_0 = I$, el método iterativo lineal haría bastantes iteraciones para $k$ pequeño, y, para $k$ suficientemente grande, ese número de iteraciones internas se reducirá a una, ya que (5.7) sería satisfecha. Así, hicimos experimentos usando problemas clásicos (discretizaciones de la ecuación de Poisson), esperando observar ese comportamiento. Sorprendentemente, el mencionado comportamiento no fue observado en ningún caso. Al contrario, el número de iteraciones del método de Krylov, en este caso GMRES, fue creciendo con $k$, de manera que el efecto de reducir $\theta_k$ fue más fuerte que precondicionar.

Una visión superficial diagnosticaría que la interpretación del Teorema 5.3 mencionada arriba está obviamente equivocada, porque el teorema dice "Existe $k_0 \in I\!N$ tal que...", y lo que sucede es que el número de iteraciones realizado en los experimentos no habría alcanzado el $k_0$ del cual se habla. Sin embargo, esta salida crea más problemas de los que resuelve. De hecho, cualquier afirmación que se refiera a límite de una sucessión incluye implícitamente un "Existe $k_0 \in I\!N$ tal que...". Por lo tanto, si nos negamos a interpretar esas palavras en términos prácticos, estaremos cuestionando el derecho de interpretar prácticamente la convergencia de cualquier sucesión, una prohibición que condenaría a la futilidad todas las teorías de convergencia.

Sin embargo, la *Discrepancia* entre teoría y práctica en este caso existe. La *Interpretación* del hecho teórico está cuestionada. No sabemos cuál es la interpretación correcta, y no conocemos la teoría que explica lo que ocurre. El cuestionamiento de la Interpretación puede llevarmos a revisar interpretaciones consagradas de muchos otros resultados teóricos y varias contradicciones deberán ser deshechas antes de llegar a una verdad aceptable. Tenemos razones más que suficientes para incluír esta teoría en el tercer nivel de desafío.

## 6. Teorías para el Futuro

En los últimos 30 años, las teorías predominantes en Optimización Numérica giraron alrededor del concepto de convergencia. Es posible que en el próximo siglo se desarrollen teorías explicativas con marcos bastante diferentes. En esta sección haremos el ejercicio "futurológico" de intentar predecir esas tendencias.

Dentro del objetivo teórico básico de explicar y prever el comportamiento de algoritmos en problemas concretos, la teoría actual encara con bastante más profundidad *los algoritmos* que *los problemas*. Una inversión de esa tendencia será oportuna. A eso se refieren nuestros dos primeros ejercícios de previsión.

a) Teorías de Clasificación de Problemas: Borges [1974, p. 708] recuerda las "ambigüedades, redundancias y deficiencias", que un doctor Franz Kuhn atribuye a cierta enciclopedia china: "En sus remotas páginas está escrito que los animales se dividen en (a) pertenecientes al Emperador, (b) embalsamados, (c) amaestrados, (d) lechones, (e) sirenas, (f) fabulosos, (g) perros sueltos, (h) incluídos en esta clasificación, (i) que se agitan como locos, (j) innumerables, (k) dibujados con un pincel finísimo de pelo de camello, (l) etcétera, (m) que acaban de romper el jarrón, (n) que de lejos parecen moscas." En la actualidad, nuestra capacidad de clasificar coherentemente problemas está en nivel similar a la clasificación de la enciclopedia borgiana. Por el contrario, la teoría de clasificación que preconizamos para el futuro será teórica porque su objetivo esencial estará en *prever* qué algoritmo resultará más adecuado para qué problema. Buenas clasificaciones posibilitarán una organización mejor de la experimentación numérica, ya que será posible elegir muestras de problemas significativos, y dar significado estadístico a los resultados. La buena clasificación teórica es necesaria. Con el "boom" de la computación, la capacidad de hacer experimentos ha crecido, y crecerá, explosivamente. Sin buenas clasificaciones corremos el riesgo de acumular montañas de datos que no sabremos interpretar. Necesitaremos saber si mil experimentos son efetivamente mil historias diferentes, o si son el mismo experimento repetido mil veces.

b) Estructuras Ocultas: Como mencionamos en la Sección 1, es común que los resultados de un algoritmo em problemas aleatorios sean diferentes que los resultados en problemas reales que, en apariencia, tendrían la misma estructura que los primeros. Cualquier experimentador numérico ha enfrentado alguma vez, con perplejidad, este fenómeno. En nuestra opinión, en muchos problemas hay "estructuras matemáticas ocultas" , que tienen influencia sobre el comportamiento de algoritmos, y que todavía no sabemos identificar. El punto de partida para descifrar estos enigmas será profundizar los problemas donde hay divergencias entre los comportamientos en los problemas reales con respecto a los aleatorios, en vez de dejarlos a un lado como hacemos actualmente. Como siempre, un buen comienzo para resolver un problema es reconocer su existencia.

c) Regularidades secuenciales: Todo experimentador numérico que haya acompañado en la terminal de la computadora la sucesión de iterandos, valores funcionales, u otros efectos observables de una sucesión producida por un algoritmo, convergente o no, ha observado regularidades en la estructura de la sucesión en su camino hacia el límite, o hacia la divergencia. Normalmente, consideramos observable apenas la convergencia y sua velocidad, de manera que miramos con curiosidad las extrañas regularidades que se presentan, pero no las registramos, seguramente por no saber qué hacer con ellas. Sin embargo, es bastante plausible que el análisis de regularidades del proceso encierre informaciones aprovechables para la resolución del problema. Inclusive, el procesamiento de las informaciones sugeridas por las regularidades podría ser desarrollado paralelamente al desarollo normal del algoritmo, usando modernas arquitecturas computacionales. Para aprovechar las regularidades, necesitamos un marco teórico que las contenga.

d) Familias de Problemas: Es fácil entender que problemas de optimización provenientes de diferentes discretizaciones de un mismo problema en dimensión infinita forman parte de la misma familia. En general, cuando un algoritmo es formulable en dimensión infinita, su comportamiento en los problemas discretizados es el mismo. De hecho, ésta es una interpretación generalmente aceptada para la convergencia de algoritmos en dimensión infinita. Sin embargo, es muy común observar identidad de comportamiento de algoritmos em familias de problemas donde varía la dimensión pero donde no hay una vinculación canónica con un único problema de dimensión infinita. Las razones de estas identidades, cuando existen, permanecen no explicadas.

e) Comportamiento de Primeiras Iteraciones: Desde la invención del Cálculo Infinitesimal, sabemos que muchas veces podemos hablar del infinito con más seguridad que de objetos finitos. En Optimización Numérica, esto se expresa de manera dramática. A veces, sabemos casi todo sobre la convergencia asintótica de una sucesión, pero no podemos decir nada sobre la iteración 15. Paradójicamente, en muchos problemas de gran importancia práctica, como los de Reconstrucción de Imágenes, son las primeras iteraciones las que interesan, porque es deseable obtener buenas imágenes después de un moderado trabajo computacional, y porque el mal condicionamiento intrínseco de los problemas disminuye el significado práctico del límite. Es obvio que la teoría de primeras iteraciones será, al princípio, mucho más difícil que la teoría asintótica, pero nos parece que si en los últimos cincuenta años se hubiera dedicado tanto esfuerzo a las primeras iteraciones como al comportamiento en el límite, gran parte del camino ya habría sido andado.

f) Dominios de Convergencia: Cuando trabajamos en convergencia local de algoritmos para Optimizacion, imaginamos al dominio de convergencia como una bola, y cuando trabajamos en convergencia global, imaginamos que es todo el espacio. En verdad, nada más diferente de auténticas bolas que los dominios de convergencia de métodos iterativos. Al mismo tiempo, cuando hay más de una solución global, la frontera entre los dominios de convergencia de diferentes soluciones dista mucho de ser a superficie suave de nuestras fantasías. Los avances recientes en Geometría Fractal comienzan a mostrarnos los verdaderos contornos de las regiones de convergencia reales. Por ahora, parecemos estar apenas contemplando interesantes curiosidades. Sin embargo, no es difícil conjeturar que después de esta etapa sabremos usar las nuevas geometrías para extraer ideas útiles para el desarrollo de algoritmos y la resolución de problemas.

## Referencias


Averick, B.M.; Carter, R.G.; Moré, J.J. [1991]: The MINPACK - 2 Test Problem Collection (Preliminary Version), Technical Memorandum No. 150, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, Illinois.

Borges, J.L. [1974]: Otras Inquisiciones, en *Obras Completas*, Emecé, Buenos Aires.

Broyden, C.G. [1965]: A class of methods for solving nonlinear simultaneous equations, *Mathematics of Computation* 19, pp. 577-593.

Broyden, C.G.; Dennis Jr., J.E.; Moré, J.J. [1973]: On the local and superlinear convergence of quasi-Newton methods, *Journal of the Institute of Mathematics and its Applications* 12, pp. 223-245.

Conn, A.R.; Gould, N.M.; Toint, Ph.L. [1992]: Numerical Experiments with the LANCELOT package (Release A) for Large - Scale Nonlinear Optimization, Report 92/16, Department of Mathematics, Facultés Universitaires ND de la Paix, Namur, Bélgica.

Dembo, R.S.; Eisenstat, S.C.; Steihaug, T. [1982]: Inexact Newton methods, *SIAM Journal on Numerical Analysis* 19, pp. 400–408.

Dennis, J.E., jr; Echebest, N.; Guardarucci, M.T.; Martínez, J.M.; Scolnik, H.D.; Vaccino, C. [1991]: A curvilinear search using tridiagonal secant updates for unconstrained optimization, *SIAM Journal on Optimization* 1, pp. 352-372.

Dennis Jr., J.E.; Moré, J.J. [1977]: Quasi-Newton methods, motivation and theory, *SIAM Review* 19, pp. 46-89.

**Dennis Jr.,J.E.; Schnabel,R.B.** [1983]:*Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice-Hall, Englewood Cliffs.

**Dennis Jr., J.E. ; Walker, H.F.** [1981]: Convergence theorems for least-change secant update methods, *SIAM Journal on Numerical Analysis* 18, pp. 949-987.

**Gay, D.M.** [1985]: Electronic mail distribution of linear programming test problems, *Mathematical Programming Committee on Algorithms Newsletter* 13, pp. 10 - 12.

**Golub, G.H.; Van Loan, Ch.F.** [1989]: *Matrix Computations*, The Johns Hopkins University Press, Baltimore MD.

**Hestenes, M.R.; Stiefel, E.** [1952]: Methods of conjugate gradients for solving linear systems, *Journal of Research of the National Bureau of Standards - Section B* 49, pp. 409 - 436.

**Gomes–Ruggiero, M.A.; Martínez, J.M.** [1992]: The Column–Updating Method for solving nonlinear equations in Hilbert space, *Mathematical Modelling and Numerical Analysis* 26, pp 309-330.

**Gomes–Ruggiero, M.A.; Martínez, J.M.; Moretti, A.C.** [1992]: Comparing algorithms for solving sparse nonlinear systems of equations, *SIAM J. on Scientific and Statistical Computing* 13, pp. 459 - 483.

**Hock, W.; Schittkowski, K.** [1981]: *Test Examples for Nonlinear Programming Codes*, Lecture Notes in Economics and Mathematical Systems Volume 187, Springer - Verlag.

**Martínez, J.M.** [1984]: A quasi–Newton method with modification of one column per iteration, *Computing* 33, pp. 353–362.

**Martínez, J.M.** [1990]: Local convergence theory of inexact Newton methods based on structured least change updates, *Mathematics of Computation* 55, pp. 143-168.

**Martínez, J.M.** [1992a]: On the Global Convergence of a Curvilinear Search Method for Unconstrained Optimization, por aparecer en *Matemática Aplicada e Computacional.*

**Martínez, J.M.** [1992b]: A Theory of Secant Preconditioners, por aparecer en *Mathematics of Computation.*

Martínez, J.M. [1992c]: On the Convergence of the Column-Updating Method, por aparecer en *Matemática Aplicada e Computacional.*

Martínez, J.M. [1992d]: On the relation between two local convergence theories of least change secant update methods, por aparecer en *Mathematics of Computation.*

Martínez, J.M.; Zambaldi, M.C. [1991]: An inverse Column-Updating Method for solving Large-Scale Nonlinear Systems of Equations, por aparecer en *Optimization Methods and Software.*

Moré, J.J. [1989]: A collection of nonlinear model problems, Preprint MCS - P60 - 0289, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, Illinois.

Moré, J.J.; Garbow, B.S. ; Hillstrom, K.E. [1981]: Testing Unconstrained Minimization Software, *ACM Transactions on Mathematical Software* 7, pp. 17–41.

Saad, Y. and Shultz, M.H. [1986]: GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems, *SIAM Journal on Numerical Analysis* 7, pp. 856–869.

# RELATÓRIOS DE PESQUISA — 1992

01/92 **Uniform Approximation the: Non-locally Convex Case** — *João B. Prolla.*

02/92 **Compactificação de $L^{\tau}_{\omega\omega}(Q)$ com $\tau$ Finito** — *A. M. Sette* and *J. C. Cifuentes.*

03/92 **Um Modelo para Aquisição da Especificação** — *Cecilia Inés Sosa Arias* and *Ariadne Carvalho.*

04/92 **Convergence Estimates for the Wavelet Galerkin Method** — *Sônia M. Gomes* and *Elsa Cortina.*

05/92 **Optimal Chemotherapy: A Case Study with Drug Resistance, Saturation Effect and Toxicity** — *M. I. S. Costa, J. L. Boldrini* and *R. C. Bassanezi.*

06/92 **On the Paper "Cauchy Completeness of Elementary Logic" of D. Mundici and A. M. Sette** — *J. C. Cifuentes.*

07/92 **What is the EM Algorithm for Maximum Likelihood Estimation in PET and How to Accelerate it** — *Alvaro R. De Pierro.*

08/92 **Bifurcation from infinity and multiple solutions for an elliptic system** — *Raffaele Chiappinelli* and *Djairo G. de Figueiredo.*

09/92 **Approximation Processes for Vector-Valued Continuous Functions** — *João B. Prolla.*

10/92 **Aplicação do Método de Fraissé à Compactificação de Lógicas com Quantificadores Co-filtro** — *A. M. Sette* and *J. C. Cifuentes.*

11/92 **Absolutely Summing Holomorphic Mappings** — *Mário C. Matos.*

12/92 **The Feynman-Dyson Proof of Maxwell Equations and Magnetic Monopoles** — *Adolfo A. Jr.* and *Waldyr A. R. Jr.*

13/92 **A Generalized Dirac's Quantization Condition for Phenomenological Non-abelian Magnetic Monopoles** — *Adolfo M. Jr.* and *Waldyr A. R. Jr.*

14/92 **Multiplicity Results for the 1-Dimensional Generalized $p$-Laplacian** — *Pedro Ubilla.*

15/92 **Nowhere Vanishing Torsion Closed Curves Always Hide Twice** — *Sueli R. Costa.* and *Maria Del Carmen R. Fuster.*

16/92 **Uniform Approximation of Continuous Convex-Cone-Valued Functions** — *João B. Prolla.*

17/92 **Monotonically Dominated Operators on Convex Cones** — *A. O. Chiacchio, J. B. Prolla, M. L. B. Queiroz* and *M. S. M. Roversi.*

18/92 **Testing the Concept of a Photon as an Extended Object in a Variation of Franson's Experiment** — *V. Buonomano, A. J. R. Madureira* and *L. C. B. Ryff.*

19/92 **A New Trust Region Algorithm for Boun Constrained Minimization** — *Ana Friedlander, José Mario Martínez* and *Sandra A. Santos.*

20/92 **A Priori Estimates for Positive Solutions of Semilinear Elliptic Systems Via Hardy-Sobolev Inequalities** — *Ph. Clément, D.G. de Figueiredo* and *E. Mitidieri.*

21/92 **On the Resolution of Linearly Constrained Convex Minimization Problems** — *Ana Friedlander, José Mario Martínez* and *Sandra A. Santos.*

22/92 **Convergência no Espectro Real de um Anel** — *J. C. Cifuentes.*

23/92 **Parallel Methods for Large Least Norm Problems** — *Alvaro R. De Pierro.* and *Jose M. Lopes.*

24/92 **A Generalization of Dirac Non–Linear Electrodynamics, and Spinning Charged Particles** — *Waldyr A. Rodrigues Jr., Jayme Vaz Jr.* and *Erasmo Recami.*

25/92 **Bifurcation of Certain Singularities of Sliding Vector Fields** — *Marco Antonio Teixeira.*

26/92 **Density of Infimum-Stable Convex Cone** — *João B. Prolla.*

27/92 **An Extended Decomposition Through Formalization in Product Spaces** — *Alvaro R. De Pierro.*

28/92 **Accelerating Iterative Algorithms for Symmetric Linear Complementarity Problems** — *Alvaro R. De Pierro* and *José Marcos Lópes.*

29/92 **Free Maxwell Equations, Dirac Equation, and Non-Dispersive de Broglie Wave–packets** — *Waldyr A. Rodrigues Jr., Jayme Vaz Jr.* and *Erasmo Recami.*

30/92 **New Results on the Equivalence of Regularization and Truncated Iteration for Ill-posed Problems** — *Reginaldo J. Santos.*

31/92 **Topological Defects: A Distribution Theory Approach** — *Patricio S. Letelier* and *Anzhong Wang.*

32/92 **Some Applications of the Generalized Khintchine's Inequality** — *Mário C. Matos.*

33/92 **A Technique for Finding The Center of A Polytope** — *Antonio Carlos Moretti.*

34/92 **On The Convergence of the Column-Updating Method** — *José Mario Martínez.*

35/92 **Post-Linear Formalism For Gravitating Strings: Crossed Straight Strings Collision** — *D. V. Gal'tsov, Yu. V. Grats* and *P. S. Letelier.*

36/92 **A Simple Model of Life Expectancy With Subjective Parameters** — *R. C. Bassanezi* and *L. C. Barros.*

37/92 **On the Growth of Caiman Crocodilus Yacare: A Strategy for Optimal Farming** — *Alejandro B. Engel* and *R. C. Bassanezi.*

38/92 **The Dirac Operator and the Structure of Riemann-Cartan-Weyl Spaces** — *Quintino A. G. de Souza* and *Waldyr A. Rodrigues, Jr.*

39/92 **On the Nature of the Gravitational Field** — *Waldyr A. Rodrigues, Jr.* and *Quintino A. G. de Souza.*

40/92 **Spinning Strings and Cosmic Dislocations** — *D. V. Gal'tsov* and *P. S. Letelier.*

41/92 **Natural Ultrabornological, Non-Complete, Normed Function Spaces** — *Antonio Gilioli* (edited by *Klaus Floret* and *Chaim S. Hönig*).

42/92 **Parallel and Semi-Parallel Immersions Into Space Forms** — *Francesco Mercuri.*

**43/92** **Periodic Motions of Sucker Rod Pumping Systems** — *Aloisio Freiria Neves* and *José Luiz Boldrini.*

**44/92** **Algorithms for Solving Nonlinear Systems of Equations** — *José Mario Martínez.*

**45/92** **Spin and Electron Structure** — *Matej Pavšič, Erasmo Recami, Waldyr A. Rodrigues Jr.* and *G. Daniele Maccarrone, Fabio Raciti, Giovanni Salesi.*

**46/92** **The Initial Value Problem For the Boussinesq Equations in a Time-Dependt Domain** — *Carlos Conca Rozende* and *Marko Rojas Medar.*

**47/92** **A New Expansion for Jacobi Function** — *E. Capelas de Oliveira.*

**48/92** **On the Equivalence of Convergences of Fuzzy Sets** — *Marko Rojas Medar* and *Heriberto Román Flores.*

**49/92** **Monotonicity and Symmetry of Solutions of Elliptic Systems in General Domains** — *Djairo G. de Figueiredo.*

**50/92** **Tendências em Programação Não Linear de Grande Porte** — *José Mario Martínez.*