

1. INTRODUCCIÓN

Sea $f: \mathbb{R}^n \rightarrow \mathbb{R}$ una función de clase $C^2(\mathbb{R}^n)$. Supongamos que $\mathbf{x}^* \in \mathbb{R}^n$ es un mínimo local estricto de f y que $\nabla^2 f(\mathbf{x}^*)$ es simétrica y positiva definida.

El método general de direcciones de descenso para aproximar un mínimo local de tal estilo es:

1.1. Método de direcciones de descenso.

Dados, la función f y $\mathbf{x}^0 \in B_R(\mathbf{x}^*) \setminus \{\mathbf{x}^*\}$ tal que $\nabla f(\mathbf{x}^0) \neq \mathbf{0}$. Si disponible, también se puede pasar el gradiente ∇f y la Hessiana $\nabla^2 f$ para evitar sus aproximaciones.

Empezamos, con $k := 0$.

MIENTRAS $\|F(\mathbf{x}^k)\| > tol$ y $k < maxIterations$

P1. Escoger un vector $\mathbf{d}^k \in \mathbb{R}^n$ de descenso ($\nabla f(\mathbf{x}^k)^T \mathbf{d}^k < 0$).

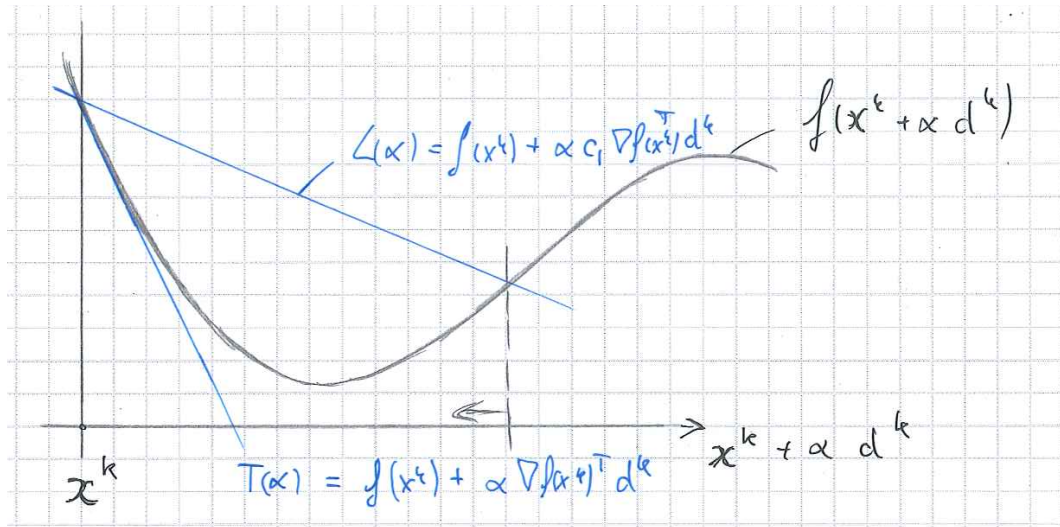
P2. Encontrar $\alpha^k \in (0, 1]$ tal que

$$f(\mathbf{x}^k + \alpha^k \mathbf{d}^k) \leq f(\mathbf{x}^k) + \alpha^k \left(c_1 \nabla f(\mathbf{x}^k)^T \mathbf{d}^k \right),$$

donde $0 < c_1 < 1$.

P3. Actualizar $\mathbf{x}^{k+1} := \mathbf{x}^k + \alpha^k \mathbf{d}^k$, redefinir $k := k + 1$ y continuar con el ciclo.

FIN

1.2. ¿Como encontrar a un α_k ? Por ahora, usen bisección. El dibujo muestra que funciona.

1.3. Direcciones de descenso.

- *Descenso por coordenadas.*

En el paso **P1** escogemos al vector $\mathbf{d}^k \in \{\pm \mathbf{e}_i\}$ tal que \mathbf{e}_i es el i -ésimo vector canónico que satisface $\nabla f(\mathbf{x}^k)^T \mathbf{d}^k = -\|\nabla f(\mathbf{x}^k)\|_\infty$. Con esta dirección se cambia solo una coordenada cuando $\mathbf{x}^k \mapsto \mathbf{x}^{k+1}$, es decir, minimizamos en una coordenada.

- *Máximo descenso.*

En cada iteración se usa el vector $\mathbf{d}^k := -\nabla f(\mathbf{x}^k) / \|\nabla f(\mathbf{x}^k)\|_2$.

Nótese, que $\|\nabla f(\mathbf{x}^k)\|$ es el criterio de terminar la iteración. (Se puede ahorrar trabajo.)

- *Dirección de Newton.*

En el caso que $\nabla^2 f(\mathbf{x}^k)$ es (simétrica) positiva definida en cada iteración, podemos escoger la dirección \mathbf{d}^k que es solución del sistema

$$\nabla^2 f(\mathbf{x}^k) \mathbf{d}^k = -\nabla f(\mathbf{x}^k).$$

Esa es una dirección de descenso, por lo que vimos en clase 3.

2. LABORATORIO

2.1. Funciones en MatLab. Escriba las funciones

```
function [xf, iter] = desCoor( f, x0, tol, maxiter )
% Purpose: descends in one variable per iteration
% In : f      ... function to minimize
%      x0      ... initial point
%      tol      ... tolerance
%      maxiter ... upper bound for iterations
%
% Out: xf      ... final approximation of x*
%      iter ... number of iterations used
```

```
function [xf, iter] = desMax( f, x0, tol, maxiter )
% Purpose: descends in largest descend direction.
```

```
function [xf, iter] = desNewton( f, x0, tol, maxiter )
% Purpose: descends in Newton direction.
```

En paso **P2** del algoritmo, se busca un $\alpha_k \in (0, 1]$ con el cual se cumple la desigualdad. A este α^k lo buscamos usando bisección empezando en $\alpha_k = 1$.

En los ejemplos (abajo) se usaron los parámetros $c_1 \stackrel{\text{def}}{=} 1/10$ y $\text{tol} \stackrel{\text{def}}{=} 10^{-5}$.

2.2. Experimentos. Usamos las funciones que hemos escrito en Laboratorio 1.

2.2.1. *La función cuadrática.* Sea f la función definida por $f(\mathbf{x}) \stackrel{\text{def}}{=} \frac{1}{2} \mathbf{x}^T A \mathbf{x} + \mathbf{b}^T \mathbf{x} + 1$, con

$$A \stackrel{\text{def}}{=} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 10 \\ 1 & 4 & 10 & 20 \end{pmatrix}, \quad \mathbf{b} \stackrel{\text{def}}{=} \begin{pmatrix} -1 \\ -1 \\ -1 \\ -1 \end{pmatrix}.$$

La función f tiene el único mínimo en $\mathbf{x}^* = (1, 0, 0, 0)^T$.

Con el punto inicial $\mathbf{x}^0 \stackrel{\text{def}}{=} (4, 4, 4, 4)^T$ se llega a la siguiente tabla.

método	iteraciones 1	iteraciones 2	Convergió
desCoor	1654	1403	Si
desMax	553	553	Si
desNewton	1	1	Si

Las “iteraciones 2” se obtuvieron con el gradiente y la Hessiana exacta.

2.2.2. *La función de Rosenbrock.* Sea $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ la función

$$f(\mathbf{x}) \stackrel{\text{def}}{=} 100(x_1^2 - x_2)^2 + (x_1 - 1)^2.$$

La función f tiene el único mínimo en $\mathbf{x}^* = (1, 1)^T$.

Con el punto inicial $\mathbf{x}^0 \stackrel{\text{def}}{=} (2, 3)^T$ se llega a la siguiente tabla.

método	iteraciones 1	iteraciones 2	Convergió
desCoor	1000	1000	No
desMax	1000	1000	No
desNewton	14	14	si

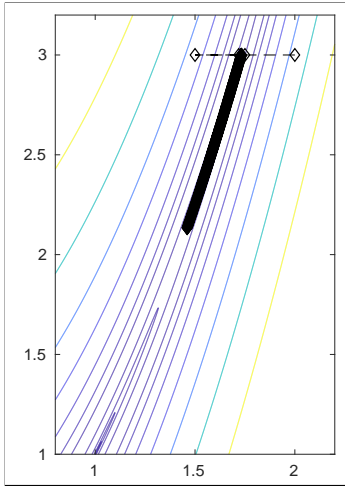
Las “iteraciones 2” se obtuvieron con el gradiente y la Hessiana exacta.

2.2.3. Visualizar lo que pasa.

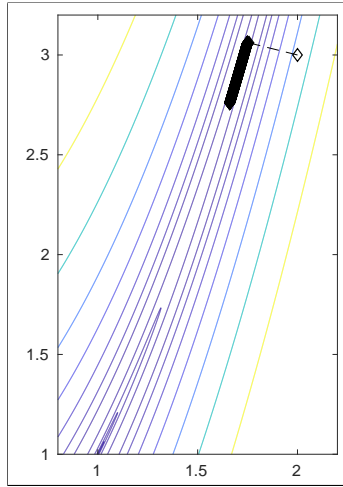
Hagan un *plot* que muestra algunos conjuntos de nivel y los primeros 1000 iteraciones.

Ejemplo en comunidad.itam: `scrLevelSet.m`

desCoor



desMax



desNewton

