ANÁLISIS APLICADO
LAB. 2

1. Introducción

Sea $f\colon \mathbb{R}^n \to \mathbb{R}$ una función de clase $C^2(\mathbb{R}^n)$. Supongamos que $\boldsymbol{x}^\star \in \mathbb{R}^n$ es un mínimo local estricto de f y que $\nabla^2 f(\boldsymbol{x}^\star)$ 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 $\boldsymbol{x}^0 \in B_R(\boldsymbol{x}^\star) \setminus \{\boldsymbol{x}^\star\}$ tal que $\nabla f(\boldsymbol{x}^0) \neq \boldsymbol{0}$. Si disponible, también se puede pasar el gradiente ∇f y la Hessiana $\nabla^2 f$ para evitar sus aproximaciones.

Empezamos, con $k \coloneqq 0$.

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

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

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

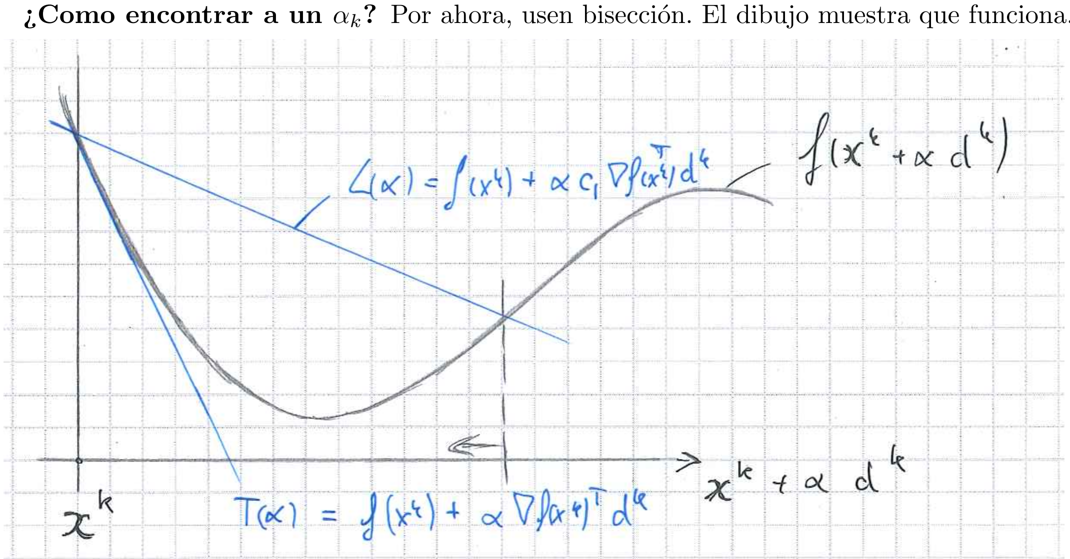
$$f(\boldsymbol{x}^k + \alpha^k \boldsymbol{d}^k) \leq f(\boldsymbol{x}^k) + \alpha^k \Big(c_1 \nabla f(\boldsymbol{x}^k)^T \boldsymbol{d}^k \Big),$$

donde $0 < c_1 < 1$.

P3. Actualizar $\boldsymbol{x}^{k+1} \coloneqq \boldsymbol{x}^k + \alpha^k \boldsymbol{d}^k$, redefinir $k \coloneqq k+1$ 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 $\boldsymbol{d}^k \in \{\pm \boldsymbol{e}_i\}$ tal que $\boldsymbol{e}_i$ es el i-ésimo vector canónico que satisface $\nabla f(\boldsymbol{x}^k)^T \boldsymbol{d}^k = -\left\|\nabla f(\boldsymbol{x}^k)\right\|_\infty$. Con esta dirección se cambia solo una coordenada cuando $\boldsymbol{x}^k \mapsto \boldsymbol{x}^{k+1}$, es decir, minimizamos en una coordenada.

- *Máximo descenso.*
  En cada iteración se usa el vector $\boldsymbol{d}^k := -\nabla f(\boldsymbol{x}^k)/\left\|\nabla f(\boldsymbol{x}^k)\right\|_2$.
  Nótese, que $\left\|\nabla f(\boldsymbol{x}^k)\right\|$ es el criterio de terminar la iteración. (Se puede ahorrar trabajo.)

- *Dirección de Newton.*
  En el caso que $\nabla^2 f(\boldsymbol{x}^k)$ es (simétrica) positiva definida en cada iteración, podemos escoger la dirección $\boldsymbol{d}^k$ que es solución del sistema
  $$\nabla^2 f(\boldsymbol{x}^k)\boldsymbol{d}^k = -\nabla f(\boldsymbol{x}^k)\,.$$
  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     ... tolarance
%      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 $c_1 \overset{\text{def}}{=\!=} 1/10$ y $\text{tol} \overset{\text{def}}{=\!=} 10^{-5}$.

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

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

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

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

Con el punto inicial $\boldsymbol{x}^0 \stackrel{\text{def}}{=\!=} (4,4,4,4)^T$ 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 $f\colon \mathbb{R}^2 \to \mathbb{R}$ la función

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

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

Con el punto inicial $\boldsymbol{x}^0 \stackrel{\text{def}}{=\!=} (2,3)^T$ 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 30 iteraciones.

Ejemplo en comunidad.itam: `scrLevelSet.m`

desCoor desMax desNewton