ANÁLISIS APLICADO
LAB. 3

1. Introducción

Vimos en Lab. 2 (clase 4) el método general de direcciones de descenso.

En clase 5, vimos que la demostración de convergencia requiere que los pasos α_k tienen que satisfacer una segunda condición (de curvatura). Lo importante de la prueba es que convergencia no requiere que escogemos pasos óptimos. Después, vimos definiciones de velocidades de convergencia (lineal, super lineal, cuadrática).

Motivos de este laboratorio:

- Vemos un algoritmo para escoger un α_k.
- Para funciones cuadráticas siempre se puede determinar el paso óptimo $\alpha_\star$ (analíticamente).
- Para funciones cuadráticas, determinar el paso óptimo $\alpha_\star$ por interpolación cuadrática.
- Verificar la velocidad de convergencia de búsqueda en linea con paso óptimo para funciones cuadráticas.

1.1. Funciones cuadráticas.

Definicion 1. *A una función $f\colon \mathbb{R}^n \to \mathbb{R}$ es cuadrática si es de la forma*

$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T Q\boldsymbol{x} + \boldsymbol{b}^T\boldsymbol{x} + c$$

donde $Q \in \mathbb{R}^{n\times n}$, $\boldsymbol{b} \in \mathbb{R}^n$, $c \in \mathbb{R}$.

1. Supón que f es una función cuadrática con una matriz simétrica positiva definida $Q \in \mathbb{R}^{n\times n}$. Demuestre que $\phi(\alpha) \overset{\text{def}}{=\!=} f(\boldsymbol{x} + \alpha\boldsymbol{d})$ es un polinomio cuadrático en α que siempre tiene un mínimo en el parámetro $\alpha_\star$. Además, si $\boldsymbol{d}$ es dirección de descenso, entonces el mínimo de ϕ esta en $\alpha_\star > 0$.

Resultado: $\alpha_\star = \frac{-\boldsymbol{d}^T\nabla f(\boldsymbol{x})}{\boldsymbol{d}^T Q\boldsymbol{d}}$

Recomendación: En la implementación usar $\frac{99}{100}\cdot\alpha_\star$. Por errores de redondeo no podemos quedar exactamente en el conjunto de nivel donde toca $\phi(\alpha_\star)$.

2. Para reducir errores en la evaluación de una forma cuadrática, podemos usar su versión simétrica. Supón que $f(\boldsymbol{x}) = \boldsymbol{x}^T Q\boldsymbol{x}$ donde $Q \in \mathbb{R}^{n\times n}$ y demuestre que

$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T(Q^T + Q)\boldsymbol{x}\,.$$

Ayuda: Considerar $Q = \frac{1}{2}(Q + Q^T) + \frac{1}{2}(Q - Q^T)$.

1.2. ¿Como encontrar a un α_k? Ya con más conocimientos.

- Para funciones cuadráticas, el ejercicio **1.** nos da una formula para el parámetro óptimo $\alpha_\star$. Haga una versión de la función `desMax`, llamada `mDesMaxCuad.m` que realiza el método de descenso con los pasos óptimos. (Debes pasar los argumentos necesarios para poder calcular esos pasos.)

- Ahora, supón que hacemos una versión de `mDesMaxCuad.m` sin los argumentos adicionales para funciones cuadráticas. Aquí, el propósito es usar los valores $\phi(0),\ \phi'(0),\ \phi(1)$, para construir un polinomio cuadrático que interpola esos datos y encontrar el mínimo de este. En aritmética de precisión infinita eso es el parámetro óptimo $\alpha_\star$.

  Encuentre la formula para este $\alpha_\star$ usando el polinomio cuadrático. (No hace falta mostrar que es el mismo valor). Pero puedes verificar que este valor coincide (ignorando errores de redondeo) con el valor exacto (apartado 1).

- Para funciones que no son cuadráticas, la idea es, primero encontrar un $\alpha_0 \leq \alpha^+$ que satisface la condición de descenso mínimo (W1) con $c_1 = 0.1$, ver clase 5 (*e.g.* con bisección). Y después usar interpolación cuadrática (o splines cúbicos) para encontrar un mínimo local en la dirección $\boldsymbol{d}^k$ en el intervalo $(0, \alpha_0)$.

  Por ejemplo: Usar interpolación cuadrática con los valores: $\phi(0),\ \phi'(0),\ \phi(\alpha_0)$ o con $\phi(0),\ \phi(\alpha_0/2),\ \phi(\alpha_0)$. La función de Rosenbrock no es cuadrática, pero sus conjuntos de nivel (en 2d) sugieren que es convexa en las direcciones de mayor descenso.

  Se debe verificar si el α encontrado satisface la condición de curvatura (W2).

- Un algoritmo que garantiza que el parámetro α encontrado satisface las condiciones (W1) y (W2) se encuentra en [Nocedal, Algorithm 3.5, 3.6]. Para entenderlo se puede hacer un *flow-diagram.*

2. Laboratorio

2.1. Funciones en MatLab. Escriba las funciones

```
function [xf, iter] = desMaxCuad( f, x0, tol, maxiter, Df, Q )
% Purpose: descends in largest descend direction with optimal step (for quadratic functions)
% In : f       ... function to minimize
%      x0      ... initial point
%      tol     ... tolarance
%      maxiter ... upper bound for iterations
%      Df      ... (optional function handle) for exact gradient of f
%      Q       ... (optional matrix) for exact hessian of f
%
% Out: xf   ... final approximation of x*
%      iter ... number of iterations used
```

```
function [xf, iter] = desMaxOpt( f, x0, tol, maxiter )
% Purpose: descends in largest descend direction.
%          For non-quadratic functions (quadratic interpolation of optimal step).
```

En los ejemplos (abajo) se usaron los parámetros $c_1 \overset{\text{def}}{=\!=} 1/10$, $c_2 \overset{\text{def}}{=\!=} 9/10$ y $\text{tol} \overset{\text{def}}{=\!=} 10^{-5}$.

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

2.2.1. Verificación. Dado una función cuadrática con una matriz Q simétrica positiva definida con eigenvalores $0 < \lambda_1 \leq \cdots \leq \lambda_n$, se puede mostrar

$$\|\boldsymbol{x} + \alpha_\star \boldsymbol{d} - \boldsymbol{x}_\star\|_Q^2 \leq \left(\frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1}\right)^2 \|\boldsymbol{x} - \boldsymbol{x}_\star\|_Q^2 \tag{1}$$

donde $\|\boldsymbol{x} - \boldsymbol{x}_\star\|_Q^2 \overset{\text{def}}{=\!=} (\boldsymbol{x} - \boldsymbol{x}_\star)^T Q (\boldsymbol{x} - \boldsymbol{x}_\star)$, $\boldsymbol{d} = -\nabla f(\boldsymbol{x})$ y $\alpha_\star$ es el parámetro óptimo de descenso.

Info: La prueba se basa en la identidad de Tarea 3 (5d) y la desigualdad de Kantorovich*. La cota muestra convergencia lineal cuando se usa el paso óptimo y la dirección de mayor descenso.*

Use tus funciones `mDesMaxCuad.m` y la función `fPascal` (del Lab. 1) para medir las tasas de convergencia. ¿Llegas en menos iteraciones que en el Lab. 2?

Good practice: Do not change the implementation. Better:

Write a script containing a loop. Inside the loop, call the method multiple times with $\text{maxIter} = 1$ and $tol = 0$ to get the next approximation, that is, `xnew = desMaxOpt(f, A, xold, tol, maxIter)` and calculate `norm(xnew- xopt)/norm(xold-xopt)`. Attention, you have to normalise the dominant eigenvector such that its euclidean norm equals one.

The same idea can be used to compute a sequence of points, and plot the points in the contour plot, *cf.*, figures in Lab. 2.

Resultados: La función cuadrática.

Sea f la función cuadrática definida en Lab. 1.

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 (de iteraciones).

método	iteraciones 1	iteraciones 2	iteraciones 3	Convergió
desMaxQuad	?	95	87	si

Las “iteraciones 1” se obtuvieron con interpolación cuadrática.

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

Las “iteraciones 3” se obtuvieron con el gradiente y la Hessiana exacta.

2.2.2. Verificación (en la función de Rosenbrock).

Mide las tasas de convergencia del método `desNewton.m` (implementado en Lab. 2).

Mide las tasas de convergencia del método `desMaxOpt.m` (con interpolación).

Resultados: La función de Rosenbrock.

Sea $f\colon \mathbb{R}^2 \to \mathbb{R}$ la función de Rosenbrock (2d), ver Lab. 1.

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 (de iteraciones).

método	iteraciones 1	Convergió
desMaxOpt	?	?
desNewton	14	si

Las iteraciones se obtuvieron con interpolación cuadrática.

23 iteraciones de desMaxCuad (usando el Hessiano (exacto) como Q)