Examen Final Análisis Aplicado

Santiago Ruiz
174667

December 12, 2020

1 Gradiente Conjugado

1. Demuestre que si los vectores no nulos $p_1, p_2, ..., p_l$ satisfacen que :

$$p_i^T A p_j = 0, \forall i \neq j,$$

y A es simétrica y positiva definida, entonces los vectores son linealmente independientes.

2. Dado este resultado, ¿Por qué el gradiente conjugado converge en a lo más n iteraciones?.

2 Quasi-Newton

1. Muestre que la segunda condición fuerte de Wolfe implica la condición de curvatura.

$$s_k^T y_k > 0.$$

2. Verifique que B_{k+1} y H_{k+1} son inversas una de la otra.

1.1 $\{p_1, \dots, p_n\}$ y $A \in M_n(\mathbb{R})$ simétrica def. pos

$\ni p_i^T A p_j = 0 \quad \forall i \neq j$

pd $\sum_{i=1}^{n} \alpha_i p_i = 0 \iff \alpha_i = 0 \ \forall i$

Supongamos que $\exists \alpha_j \neq 0$ y $\sum_{i=1}^{n} \alpha_i p_i = 0$

con $j \in \{1, \dots, n\}$ fija

$$\sum_{i=1}^{n} \alpha_i p_i = \bar{0}$$

Multiplicando A simétrica def positiva por la izquierda

$$A\sum_{i=1}^{n}\alpha_i p_i = A(\bar{0}) \Rightarrow \sum_{i=1}^{n}\alpha_i A p_i = \bar{0}$$

Multiplicando p_j^T por la izquierda

$$p_j^T\sum_{i=1}^{n}\alpha_i A p_i = p_j^T\bar{0} \Rightarrow \sum_{i=1}^{n}\alpha_i p_j^T A p_i = 0$$

Como sabemos que $p_j^T A p_i = 0 \ \forall i \neq j$

$$\Rightarrow \sum_{i=1}^{n}\alpha_i p_j^T A p_i = \alpha_j p_j^T A p_j = 0$$

Como A es def pos. $\Rightarrow \forall x \in \mathbb{R}^n$

$$x^T A x > 0$$

$\therefore \alpha_j p_j^T A p_j = 0 \iff \alpha_j = 0$!

$\therefore \alpha_j$ debe ser cero para que

$\sum_{i=1}^{n}\alpha_i p_i = 0 \therefore \{p_1, \dots, p_n\}$ es li $\blacksquare$

1.2 Esto es equivalente a probar el siguiente Teorema visto en clase:

"Para todo $x_0 \in \mathbb{R}^n$, la sucesión $\{x_i\}$ dada por $x_{i+1} = \alpha_i p_i$ $\quad \alpha_i = \dfrac{-r_i^T p_i}{p_i^T A p_i}$

converge a x^* en al menos n pasos"

Como ya fue demostrado, $\{p_0, \dots, p_{n-1}\}$ son l.i $\Rightarrow$ span $\{p_1, \dots, p_n\} = \mathbb{R}^n$

$\therefore$ $x^* - x_0 \in \mathbb{R}^n$ se puede escribir como una combinación lineal de $\{p_0, \dots, p_{n-1}\}$

$$x^* - x_0 = \sum_{i=0}^{n-1} \sigma_i p_i \quad \text{p.a escalar } \sigma_i$$

Multiplicamos $p_k^T A$, $k \in \{0, \dots, n-1\}$ por la izquierda

$$p_k^T A (x^* - x_0) = p_k^T A \sum_{i=1}^{n-1} \sigma_i p_i = \sum_{i=1}^{n-1} \sigma_i p_k^T A p_i$$

pero sabemos que $p_k^T A p_i = 0$ $k = i$

$$p_k^T A (x^* - x_0) = \sigma_k p_k^T A p_k$$

Como $p_k^T A p_k > 0$ porque A es def. pos

$$\sigma_k = \frac{p_k^T A (x^* - x_0)}{p_k^T A p_k}$$

Ahora, como la sucesion $\{x_k\}$ está dada por $x_k = x_{k-1} + \alpha_{k-1} p_{k-1}$

$\Rightarrow x_k = x_0 + \alpha_0 p_0 + \dots + \alpha_{k-1} p_{k-1}$

Multiplicando $p_k^T A$ por la izquierda

$$p_k^T A x_k = p_k^T A x_0 \Rightarrow p_k^T A (x_k - x_0) = 0$$

$$\therefore p_k^T A (x^* - x_0) = p_k^T A (x^* - x_k)$$
$$= p_k^T (b - A x_k) = -p_k^T r_k$$

$$\therefore \sigma_k = \frac{-p_k^T r_k}{p_k^T A p_k} = \alpha_k$$

$\therefore \{x_i\} \longrightarrow x^*$ en k pasos, con $k \leq n$

2.1 pd $|\nabla f(x_k+\alpha_k p_k)^T p_k| \leq c_2 |\nabla f_k^T p_k|$

$\Rightarrow s_k^T y_k > 0$

Recordemos que $x_{k+1} = x_k + \alpha_k p_k$

$s_k = x_{k+1} - x_k$ y $y_k = \nabla f_{k+1} - \nabla f_k$

Además $B_{k+1} s_k = y_k$

La segunda condición de Wolfe indica que

$|\nabla f(x_k+\alpha_k p_k)^T p_k| \leq c_2 |\nabla f_k^T p_k|$

$\Rightarrow \nabla f(x_k+\alpha_k p_k)^T p_k \leq c_2 |f_k^T p_k|$

$\Rightarrow \nabla f(x_k+\alpha_k p_k)^T p_k \geq -c_2 |\nabla f(x_k)^T p_k|$

Como $\nabla f(x_k)^T p_k < 0$ por p_k dirección de descenso

$\Rightarrow -c_2 |\nabla f(x_k)^T p_k| = c_2 \nabla f(x_k)^T p_k$

$\nabla f(x_k+\alpha_k p_k)^T p_k - \nabla f(x_k)^T p_k$

$= (c_2 - 1) \nabla f(x_k)^T p_k > 0$ $\circledast$

porque $c_2 < 1 \Rightarrow c_2 - 1 < 0$ y $\nabla f(x_k)^T p_k < 0$

$$s_k^T y_k = (x_{k+1} - x_k)^T (\nabla f_{k+1} - \nabla f_k)$$

$$= (\cancel{x_k} + \alpha_k p_k - \cancel{x_k})^T (\nabla f_{k+1} - \nabla f_k)$$

$$= \alpha_k p_k^T (\nabla f_{k+1} - \nabla f_k)$$

$$= \alpha_k (p_k^T \nabla f_{k+1} - p_k^T \nabla f_k) > 0$$

porque α_k es el tamaño del paso, i. $\alpha_k > 0$ y $p_k^T \nabla f_{k+1} - p_k^T \nabla f_k > 0$ por (*)

ya que $p_k^T \nabla f_{k+1} = \nabla f_{k+1}^T p_k$ y $p_k^T \nabla f_k = \nabla f_k^T p_k$

$\therefore \underline{s_k^T y_k > 0}$

2.2 pd B_{k+1} y H_{k+1} son tal que

$$B_{k+1}^{-1} = H_{k+1}$$

Por las fórmulas de Sherman-Morrison-Woodbury, tenemos que

$$B_{k+1} = B_k - \frac{B_k y_k y_k' B_k}{y_k^T B_k y_k} + \frac{s_k s_k^T}{y_k^T s_k}$$

y ademas, sabemos que

$$H_{k+1} = (I - \rho_k y_k s_k^T) H_k (I - \rho_k s_k y_k^T) + \rho_k s_k s_k^T$$

con $\rho_k = 1/y_k^T s_k$

$$B_{k+1} H_{k+1} = B_{k+1} (I - \rho_k y_k s_k) H_k (I - \rho_k s_k y_k^T)$$
$$+ B_{k+1} \rho_k y_k y_k^T$$
$$= (B_{k+1} - \rho_k y_k y_k^T) H_k (I - \rho_k y_k s_k^T) + \rho_k y_k s_k^T$$

porque $B_{k+1} s_k = y_k$

$$= \left(B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k}\right) H_k (I - \rho_k y_k s_k^T)$$
$$+ \rho_k y_k s_k^T$$

$$= \left(I - \frac{B_k s_k s_k^T}{s_k^T B_k s_k}\right)(I - \rho_k y_k s_k^T) + \rho_k y_k s_k^T$$

$$= I \cancel{- \rho_k y_k s_k^T} - \frac{B_k s_k s_k^T}{s_k^T B_k s_k} + \frac{B_k s_k s_k^T \rho_k y_k s_k^T}{s_k^T B_k s_k}$$

$$\cancel{+ \rho_k y_k s_k^T}$$

$$= I - \frac{B_k s_k s_k^T}{s_k^T B_k s_k} + \frac{\rho_k B_k s_k \overset{\rho_k^{-1}}{\cancel{s_k^T y_k}} s_k^T}{s_k^T B_k s_k}$$

$$= I - \frac{B_k s_k s_k^T}{\cancel{s_k^T} B_k s_k} + \frac{\overset{1}{\cancel{\rho_k \rho_k^{-1}}} B_k s_k s_k^T}{s_k^T B_k s_k} \quad 0$$

$$= I$$