Esteban Martínez

Examen Final Análisis Aplicado

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. 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.

sea $n = l$

sea $\alpha_i \in \mathbb{R}$ $\forall i \mid 1 \leq i \leq n$

p.d. $\{p_1, \dots, p_n\}$ son l.i

p.d. $\alpha_1 p_1 + \alpha_2 \cdot p_2 + \dots + \alpha_n \cdot p_n = 0 \Rightarrow \alpha_i = 0 \ \forall i$

supongamos que $\{p_1, \dots, p_n\}$ no son l.i. , entonces $\exists \ \alpha_k \neq 0$ con $1 \leq k \leq n$

tal que $\alpha_1 \cdot p_1 + \dots + \alpha_k \cdot p_k + \dots + \alpha_n \cdot p_n = 0$, $\alpha_k \neq 0$

entonces tenemos que

$$A(\alpha_1 \cdot p_1 + \dots + \alpha_k \cdot p_k + \dots + \alpha_n \cdot p_n) = 0$$

$$\Rightarrow \quad p_k^T \cdot A(\alpha_1 p_1 + \dots + \alpha_k \cdot p_k + \dots + \alpha_n \cdot p_n) = 0$$

$$\Rightarrow p_k^T \cdot A \cdot \alpha_1 \cdot p_1 + \dots + p_k^T \cdot A \cdot \alpha_k \cdot p_k + \dots + p_k^T \cdot A \cdot \alpha_n \cdot p_n = 0$$

$$\Rightarrow \alpha_1 (p_k^T \cdot A \cdot p_1) + \dots + \alpha_k (p_k^T \cdot A \cdot p_k) + \dots + \alpha_n \cdot (p_k^T \cdot A \cdot p_n) = 0$$

$$\Rightarrow 0 + 0 + \dots + \alpha_k \cdot (p_k^T \cdot A \cdot p_k) + 0 + \dots + 0 = 0$$

$$\Rightarrow \quad \alpha_k (p_k^T \cdot A \cdot p_k) = 0$$

pero $\alpha_k \neq 0$

$$\Rightarrow \quad p_k^T \cdot A \cdot p_k = 0$$

pero como A es positiro definida

$$p_k^T \cdot A \cdot p_k > 0$$

$\therefore \{p_1, p_2, \dots, p_k\}$ son l.i.

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

$$(1-c_2)\, |\nabla f(x_k) \cdot p_k \leq$$

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

recordando que la iteracción entre soluciones está dada por la siguiente formula

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

$$\text{con} \quad \alpha_k = \frac{-r_k^T \cdot p_k}{p_k^T \cdot A \cdot p_k}$$

ahora bien, como arriba probamos que $\{p_i\}$ son $l.i.$

$$\text{span}\, \{p_i \mid 1 \leq i \leq n\} = \mathbb{R}^n$$

como $x_* - x_0 \in \mathbb{R}^n$ con x_* solución del problema

$$\therefore \quad x_* - x_1 = \gamma_1 p_1 + \cdots \gamma_n \cdot p_n$$

$$\text{entonces} \quad \gamma_i = \frac{p_i^T A (x_* - x_0)}{p_i^T A \cdot p_i}$$

ahora bien, regresando a cómo se construyen la iteración x_k

$$x_k = x_0 + \alpha_0 \cdot p_0 + \cdots + \alpha_{k-1} \cdot p_{k-1}$$

$\iff \quad x_k - x_0 = \alpha_1 \cdot p_1 + \cdots + \alpha_k \cdot p_k$, multiplicando por $p_k^T \cdot A$

tenemos $p_k^T \cdot A\, (x_k - x_1) = 0$

$$\text{entonces} \quad p_k^T \cdot A(x_* - x_0) = p_k^T A\, (x_* - x_k) = p_k^T \cdot r_k$$

$\therefore$ llegamos a que $\alpha_k = \gamma_k$ lo cual implica que el algoritmo termina en a lo más n iteraciones pues $x_n = x_*$ //

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

$$s_k^T y_k > 0.$$

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

la segunda condición de Wolfe dice

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

al ser p una dirección de descenso tenemos

$$\nabla f(x_k + \alpha_k \cdot p_k)^T \cdot p_k \geq c_2 \nabla f_k^T \cdot p_k \qquad \text{con } c_2 < 1$$

$$\Rightarrow \quad \nabla f(x_{k+1})^T \cdot p_k \geq c_2 \nabla f(x_k)^T \cdot p_k$$

$$\Rightarrow \quad \nabla f(x_{k+1})^T \cdot p_k - \nabla f(x_k)^T \cdot p_k \geq c_2 \nabla f(x_k)^T \cdot p_k - \nabla f(x_k)^T \cdot p_k$$

$$\Rightarrow \quad y_k^T \cdot p_k \geq (c_2 - 1)\, \alpha_k \nabla f_k^T \cdot p_k$$

de nuevo, como $c_2 < 1$, p_k dirección de descenso

tenemos

$$y_k^T \cdot p_k > 0$$

ahora bien $s_k = x_{k+1} - x_k = \alpha_k \cdot p_k$

$\therefore$ tenemos que $\dfrac{y_k^T \cdot s_k}{\alpha_k} > 0$

pero α_k sabemos que es > 0

$$\therefore \quad y_k^T \cdot s_k > 0$$

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

dado que $B_K^{-1} = H_K$

$$H_{K+1} = (I - \rho_K s_K y_K^T) \cdot H_K \cdot (I - \rho_K y_K s_K^T) + \rho_K s_K \cdot s_K^T$$

$$B_{K+1} = B_K - \frac{B_K s_K s_K^T \cdot B_K}{s_K^T \cdot B_K \cdot s_K} + \frac{y_K \cdot y_K^T}{y_K^T \cdot s_K}$$

p.d. $B_{K+1} \cdot H_{K+1} = I$

\\ voy a quitar los subíndices de y, s, ρ

$$B_{K+1} \cdot H_{K+1} = B_{K+1} (I - \rho s y^T) H_K (I - \rho y s^T) + \rho s \cdot s^T$$

$$= (B_{K+1} - \rho y \cdot y^T) \cdot H \cdot (I - \rho y s^T) + \rho \cdot y s^T$$

$$= \left(B_K - \frac{B_K \cdot s \cdot s^T B_K}{s^T \cdot B_K \cdot s}\right) \cdot H_K \left(I - \rho y \cdot s^T\right) + \rho \cdot y \cdot s^T$$

$$= \left(I - \frac{B_K s \cdot s^T}{s^T \cdot B_K \cdot s}\right) (I - \rho \cdot y \cdot s^T) + \rho \cdot y \cdot s^T$$

$$= \left(I - \rho y s^T\right) - \left(\frac{B_K s \cdot s^T}{s^T \cdot B \cdot s}\right) \left(I - \rho \cdot y \cdot s^T\right) + \rho \cdot y \cdot s^T$$

$$= I - \left(\frac{B_K s \cdot s^T}{s^T \cdot B_K \cdot s}\right) (I - \rho \cdot y \cdot s^T)$$

ahora bien, si tomamos en cuenta el hecho

$$\frac{B_K s \cdot s^T}{s^T \cdot B_K \cdot s} \cdot \rho y \cdot s^T = \frac{1}{s^T \cdot y} \frac{B_K \cdot s (s^T \cdot y) \cdot s^T}{s^T \cdot B_K s} = \frac{B \cdot s \cdot s^T}{s^T \cdot B \cdot s}$$

tenemos que

$$\left(\frac{B_K s \cdot s^T}{s^T \cdot B_K \cdot s}\right) (I - \rho \cdot y \cdot s^T) = \frac{B_K \cdot s \cdot s^T}{s^T \cdot B_K \cdot s} - \frac{B_K \cdot s \cdot s^T}{s^T \cdot B \cdot s} = 0$$

$$\therefore B_{K+1} \cdot H_{K+1} = I - 0 = I$$