Examen Final (Análisis Aplicado)

Leonardo David Plata Martínez
167480

Gradiente conjugado

1.- $(p_1, p_2, \ldots, p_l)$ satisfacen $p_i^T A p_j = 0 \quad \forall i \neq j$ y A simétrica y positiva definida, ent los vectores son LI.

Propiedad conjugada: $p_i^T A p_j = 0 \ \forall i \neq j$

P.d. A simétrica y positiva definida y los vectores son conjugados $\Rightarrow$ Los vectores son LI

Por contradicción

Suponemos que si son conjugados $\Rightarrow$ No son LI.

Como $(p_1, \ldots, p_l)$ no son linealmente independientes, podemos expresar al cero como una combinación lineal de p_i

$$\sum \alpha_i p_i = 0 \quad \text{con algun } \alpha_i \neq 0$$

$$\Updownarrow$$

$$\sum \alpha_i A p_i = 0 \iff \sum \alpha_i p_j^T A p_i = 0 \quad \text{con cualquier } p_j \text{ del cjto}$$

Por propiedad conjugada $(p_j^T A p_i = 0 \ i \neq j)$

solo sobrevive el término $\alpha_j p_j^T A p_j = 0$

Usando que A positiva definida, entonces $p_i^T A p_i > 0$, lo que nos lleva a que $\alpha_j = 0$ necesariamente. Ésto se cumple para cualquier α_j, y $\alpha_i = 0 \ \forall i$

$\therefore (p_1, \ldots, p_l)$ son LI.

1.2 ¿Por qué el gradiente conjugado converge en a lo más n iter?

Dado que $(p_0, \ldots, p_{n-1})$ son LI, tendremos que A tendrá a lo más n direcciones conjugadas que corresponden a los vectores $\{p_i\}$

1.2) ¿Por qué el gradiente conjugado converge a lo más en n iter?

Dado que $\{P_i\}$ son vectores L.I. Por el inciso anterior, sabemos que generan todo el espacio $\mathbb{R}^n$ y por lo tanto A tendría a lo más n direcciones conjugadas. Y el método de gradiente conjugado minimiza sobre cada una de las direcciones conjugadas. Es por esto que el algoritmo termina en a lo más n iteraciones.

Quasi-Newton

2.- segunda condición Fuerte de wolfe: $|\nabla f(x_k + \alpha_k P_k)^T P_k| \leq C_2 |\nabla f_k^T P_k|$

P.d. $|\nabla f(x_k + \alpha_k P_k)^T P_k| \leq C_2 |\nabla f_k^T P_k| \Rightarrow s_k^T y_k > 0$

$\to$ con $0 < C_1 < C_2 < 1$

Quitando el valor abs de la parte izq, queda

$$-C_2 |\nabla f_k^T P_k| \leq \nabla f(x_k + \alpha_k P_k)^T P_k \leq C_2 |\nabla f_k^T P_k|$$

Como P_k es una dirección de descenso, $\nabla f_k^T P_k < 0$

$$\Rightarrow C_2 \nabla f_k^T P_k \leq \nabla f(x_k + \alpha_k P_k)^T P_k$$

$$\Leftrightarrow C_2 \nabla f_k^T P_k - \nabla f_k^T P_k \leq \nabla f(x_k + \alpha_k P_k)^T P_k - \nabla f_k^T P_k$$

$$\Rightarrow \nabla f_k^T P_k (C_2 - 1) \leq \nabla f(x_k + \alpha_k P_k)^T P_k - \nabla f_k^T P_k = y_k^T P_k$$

$C_2 < 1 \to C_2 - 1 < 0 \to \nabla f_k^T P_k (C_2 - 1) > 0$

$$\Rightarrow 0 < \nabla f_k^T P_k (C_2 - 1) \leq y_k^T P_k$$

$$\Leftrightarrow 0 < \nabla f_k^T \alpha_k P_k (C_2 - 1) \leq y_k^T \underbrace{\alpha_k P_k}_{s_k}$$

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

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

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

P.d $B_{k+1} H_{k+1} = I$

Es importante notar que

$$\rho_k \frac{B_k s_k s_k^T}{s_k^T B_k s_k} y_k s_k^T = \frac{1}{s_k^T y_k} \frac{B_k s_k (s_k^T y_k) s_k^T}{s_k^T B_k s_k} = \frac{B_k s_k s_k^T}{s_k^T B_k s_k}$$

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

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

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

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

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

$$= I$$

$\therefore B_{k+1}$ y H_{k+1} son inversas entre sí