Juan Pablo Rosas Flores

165879

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 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?.

1.- Por la hipótesis asumimos que los vectores $p_1, p_2, \ldots, p_l$ son conjugados, ahora bien, supongamos que no son linealmente independientes, entonces podemos escribir

$p_i = \sum_{k \neq i} \sigma_k p_k$, es decir, escribir al i-ésimo como una comb. lineal de los demás. Ahora multiplicamos por izquierda a $p_j^T A$ para cualquier $j \neq i$, y de esta forma:

$$0 = p_j^T A p_i = p_j^T A \sum_{k \neq i} \sigma_k p_k$$

$$= \sigma_1 \cancel{p_j^T A p_1}^{\,0} + \sigma_2 \cancel{p_j^T A p_2}^{\,0} + \ldots + \sigma_j p_j^T A p_j + \ldots + \sigma_l \cancel{p_j^T A p_l}^{\,0}$$

$\Rightarrow \forall j \neq i$ tenemos que: $0 = \sigma_j p_j^T A p_j$

$* \nearrow 0$ por ser conjugados

y como A es posit. def: $p_j^T A p_j > 0 \Rightarrow \sigma_j = 0$

$\therefore$ Si $\forall j \neq i$ $\sigma_j = 0$ tenemos que

$p_i = \sum_{k \neq i} \sigma_k p_k = 0$! pues los p_i's son no-nulos.

$\therefore$ Si son l.i. pues al suponer lo contrario llegamos a un absurdo //

2.- Por el inciso anterior dado que los $\{P_i\}$ son l.i.,

$\mathbb{R}^n \subseteq \text{span}\{P_i\}$, queremos ver que $\{x_k\} \longrightarrow x^*$ en máx. n iteraciones donde $x_0 \in \mathbb{R}^n$, x^* es la sol. óptima.

$$\Rightarrow \quad x^* - x_0 = \sigma_0 P_0 + \sigma_1 P_1 + \cdots + \sigma_{n-1} P_{n-1} \quad \text{para algunos } \sigma_k$$

$$\Leftrightarrow \quad P_k^T A (x^* - x_0) = P_k^T A (\sigma_0 P_0 + \sigma_1 P_1 + \cdots + \sigma_{n-1} P_{n-1})$$

y por hipótesis $P_j^T A P_i = 0 \quad \forall i \neq j$

$$\Rightarrow \quad \sigma_k = \frac{P_k^T A (x^* - x_0)}{P_k^T A P_k}, \quad y \quad x_{k+1} = x_k + \alpha P_k$$

$$y \quad \min \phi = Ax - b = r(x) \quad \Rightarrow \quad r_k = A x_k - b$$

$$\Rightarrow \quad \alpha_k = \frac{-r_k^T P_k}{P_k^T A P_k}$$

y al usar $x_k = x_0 + \sigma_0 P_0 + \sigma_1 P_1 + \cdots + \sigma_{k-1} P_{k-1}$:

$$P_k^T A (x_k - x_0) = 0 \Rightarrow P_k^T A (x^* - x_0) = P_k^T A (x^* - x_k) = -P_k^T P_k$$

con $\sigma_k = \alpha_k$ vemos que máx. en n iteraciones se converge //

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.- La segunda condición fuerte de Wolfe nos dice que

(1) $|\nabla f(x_k + \alpha_k p_k)^T p_k| \leq c |\nabla f(x_k)^T p_k| \quad 0<c<1$

$\iff \nabla f(x_k+\alpha_k p_k)^T p_k \geq -c|\nabla f(x_k)^T p_k|$, pero como p_k es dirección de descenso:

$$\iff \nabla f(x_k+\alpha_k p_k)^T p_k = c\nabla f(x_k)^T p_k$$

(2) $\iff \nabla f(x_k+\alpha_k p_k)^T p_k - \nabla f(x_k)^T p_k = (c-1)\nabla f(x_k)^T p_k > 0$ pues $c-1<0$

Ahora bien, $y_k = \nabla f(x_k+\alpha_k p_k) - \nabla f(x_k) \wedge s_k = \alpha_k p_k$

$\therefore$ si a (2) le mult. α_k tenemos:

$$\nabla f(x_k+\alpha_k p_k)^T \alpha_k p_k - \nabla f(x_k)^T \alpha_k p_k = (c-1)\nabla f(x_k)^T \alpha_k p_k > 0$$

y al sustituir $s_k = \alpha_k p_k$:

$$\left(\nabla f(x_k+\alpha_k p_k)^T - \nabla f(x_k)^T\right) s_k = (c-1)\nabla f(x_k)^T s_k > 0$$

y al sustituir $y_k = \nabla f(x_k+\alpha_k p_k) - \nabla f(x_k)$:

$$(y_k)^T s_k > 0 \iff s_k^T y_k > 0 \quad //$$

2.- tenemos: (1) $B_{k+1} = B_k - \dfrac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \dfrac{y_k y_k^T}{y_k^T s_k} = B_k - \dfrac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \rho_k y_k y_k^T$

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

con $\rho_k = \dfrac{1}{y_k^T s_k}$

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

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

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

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

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