Análisis Aplicado, Examen Final

Oscar Alejandro Aguilar Castillo - 173718

12 de diciembre de 2020

1. Gradiente Conjugado

1.1. 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.

1. Procederemos por contradicción:

Supongamos que los vectores $p_i \neq 0$ no son linealmente indep. $\forall i$

$\Rightarrow$ Si no son l.i. podemos encontrar cts α_i no todas distintas de 0 t.q.

$$\sum_i \alpha_i p_i = 0$$

Ahora bien, si multiplicapor por una A simétrica y pos. definida:

$$\Rightarrow \quad A \sum_i \alpha_i p_i = \sum_i A \alpha_i p_i = \sum_i \alpha_i A p_i = 0$$

Posteriormente, multiplicando por p_j^T:

$$\Rightarrow \quad p_j^T \sum_i \alpha_i A p_i = \sum_i p_j^T \alpha_i A p_i = \sum_i \alpha_i (p_j^T A p_i) = 0$$

Por otro lado, por hip. sabemos que $p_j^T A p_i = 0 \ \forall i \neq j$, ahora, analizando $i = j$:
como A por hip. es pos. def $\Rightarrow p_j^T A p_i > 0$

$$\Rightarrow \text{En particular: } \frac{\alpha_j (\cancel{p_j^T A p_j})}{\cancel{p_j^T A p_j}} = 0 \iff \alpha_j = 0 \ \forall j \ ¡$$

↳ Llegamos a una contradicción $\therefore$ los valores p_i son l.i. $\forall i$.

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

Como tenemos que las direcciones p_i son linealmente independientes, logran generar todo $\mathbb{R}$. Entonces en clase vimos un teorema que mostraba que para cualquier $x_0 \in \mathbb{R}$ la sucesión x_k generada con el método de gradiente conjugado, converge a la solución x^* en a lo más n pasos.

Por otro lado, como complemento, en cuestiones de convergencia, también podemos ver que a lo más, la matriz formada por los vectores p_i linealmente independientes, tendría a lo más n eigenvalores distintos por lo tanto, usando otro teorema de convergencia, sabemos que si tenemos una matriz con n eigenvalores distintos, entonces la iteración del gradiente conjugado terminará a lo más en n iteraciones.

2. Quasi-Newton

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

$$s_k^T y_k > 0.$$

1. Tenemos que la segunda condición fuerte de Wolfe está dada por:

$$|\nabla f(x_k + \alpha_k p_k)^T p_k| \leq c_2 |\nabla f(x_k)^T p_k| \quad \text{lo cual implica:}$$

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

En particular, como p_k es una dirección de descenso, se sigue que:

$$\Rightarrow \nabla f(x_k + \alpha_k p_k)^T p_k - \nabla f(x_k)^T p_k = (c_2 - 1) \nabla f(x_k)^T p_k \overset{*}{>} 0$$

$\overbrace{\text{Como tomamos } c_2 \in (c_1, 1) \Rightarrow c_2 < 1}^{*}$ & multiplicando por α_k:

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

Redefiniendo, sea $s_k = \alpha_k p_k$ & $y_k = \nabla f(x_k + \alpha_k p_k) - \nabla f(x_k)$

$$\Rightarrow y_k^T s_k \geq (c_2 - 1) \alpha_k \nabla f_k^T p_k \quad \text{lo cual cumple las cond. de curvatura}$$

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

2.- Definimos $B_{k+1} = B_k - \left(\dfrac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} \right) + \dfrac{y_k y_k^T}{y_k^T s_k}$

y por otro lado: $H_{k+1} = (Id - \rho_k s_k y_k^T) H_k (Id - \rho_k y_k s_k^T) + \rho_k s_k s_k^T$

↳ Si en efecto B_{k+1} & H_{k+1} fueran inversas la una de la otra, se seguiría entonces que $B_{k+1} H_{k+1} = H_{k+1} B_{k+1} = B_{k+1}(B_{k+1})^{-1} = H_{k+1}(H_{k+1})^{-1} = Id$

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

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

$$= \left(B_{k+1} - \frac{1}{y_k^T s_k} (s_k y_k^T) \right) H_k (Id - \rho_k y_k s_k^T) + \rho_k y_k s_k^T =$$

$$\overset{\text{defn } B_{k+1}}{=} \left(B_k + \left(\frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} \right) (\rho_k s_k y_k^T) \right) H_k (Id - \rho_k y_k s_k^T) + \rho_k y_k s_k^T =$$

$$= \left(B_k + \frac{1}{\cancel{s_k^T y_k}} \left(\frac{B_k s_k (\cancel{s_k^T y_k}) s_k^T B_k}{s_k^T B_k s_k} \right) \right) H_k (Id - \rho_k y_k s_k^T) + \rho_k y_k s_k^T =$$

$H_k = B_k^{-1}$

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

$$= Id + \frac{B_k s_k s_k^T}{s_k^T B_k s_k} - \frac{1}{\cancel{y_k^T s_k}} \left(\frac{\cancel{y_k^T s_k} B_k s_k s_k^T}{s_k^T B_k s_k} \right) - \cancel{\rho_k y_k s_k^T} + \cancel{\rho_k y_k s_k} = Id + \cancel{\frac{B_k s_k s_k^T}{s_k^T B_k s_k}} - \cancel{\frac{B_k s_k s_k^T}{s_k^T B_k s_k}} = Id.$$

∴ En efecto B_{k+1} & H_{k+1} son inversas la una de la otra