Daniela Bugeda Reyes 174308
Examen Final Análisis Aplicado I

① Teoría:

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.

Solución:

P.d. $\{p_1, p_2, \dots, p_n\}$ son linealmente independientes
P.d. $\alpha_1 p_1 + \alpha_2 p_2 + \cdots + \alpha_n p_n = 0$ con $\alpha_i = 0 \;\; \forall i = 1, \dots, n$.

Supongamos que no y procedamos por contradicción.
Supongamos pues que el vector p_k (SPG) puede escribirse como combinación lineal de los otros $n-1$ p_i, esto es:

(1) $p_k = \beta_1 p_1 + \beta_2 p_2 + \cdots + \beta_{k-1} p_{k-1} + \beta_{k+1} p_{k+1} + \cdots + \beta_n p_n$ con al menos un $\beta_j \neq 0$.

Ahora, multiplicamos (1) ambos lados por la izquierda por $p_k^T A$ y obtenemos lo siguiente:

$$p_k^T A p_k = p_k^T A (\beta_1 p_1 + \beta_2 p_2 + \cdots + \beta_{k-1} p_{k-1} + \beta_{k+1} p_{k+1} + \cdots + \beta_n p_n)$$

$$\Rightarrow p_k^T A p_k = p_k^T A \beta_1 p_1 + p_k^T A \beta_2 p_2 + \cdots + p_k^T A \beta_{k-1} p_{k-1} + p_k^T A \beta_{k+1} p_{k+1} + \cdots + p_k^T A \beta_n p_n$$

como β_j es un escalar $\forall j$, podemos "sacarlo" de cada multiplicación:

$$\Rightarrow p_k^T A p_k = \beta_1 (p_k^T A p_1) + \beta_2 (p_k^T A p_2) + \cdots + \beta_{k-1} (p_k^T A p_{k-1}) + \beta_{k+1} (p_k^T A p_{k+1}) + \beta_n (p_k^T A p_n)$$

Por hipótesis tenemos que $p_i^T A p_j = 0 \;\; \forall i \neq j$

$$\Rightarrow p_k^T A p_k = \beta_1 (0) + \beta_2 (0) + \cdots + \beta_{k-1}(0) + \beta_{k+1}(0) + \cdots + \beta_n (0)$$

$\therefore p_k^T A p_k = 0$ ¡! pues p_k es no nulo y por hipótesis tenemos que $p_k^T A p_k > 0$ pues A positiva definida

$\therefore p_k$ no puede expresarse como combinación lineal de $\{p_i\}$
$\therefore \{p_1, \dots, p_n\}$ es linealmente independiente $\blacksquare$

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

Solución:

→ Teorema 5.1 (Nocedal)

En clase probamos que $\forall x_0 \in \mathbb{R}^n$ la secuencia $\{x_k\}$ generada por el algoritmo de direcciones conjugadas converge a la solución x^* del sistema lineal $Ax = b$ en a lo más n iteraciones.

A grandes rasgos, la razón de eso es que, dado que las direcciones conjugadas $\{p_i\}$ son linealmente independientes, entonces podemos expresar $x^* - x_0$ como combinación lineal de p_i's:

$$x^* - x_0 = \sigma_0 p_0 + \cdots + \sigma_{n-1} p_{n-1} \quad \text{para ciertos } \sigma_k$$

Haciendo álgebra llegaríamos a $\sigma_k = \dfrac{p_k^T A (x^* - x_0)}{p_k^T A p_k}$

Como x_k es generada por el algoritmo de direcciones conjugadas por hipótesis, entonces:

$$x_k = x_0 + \alpha_0 p_0 + \alpha_1 p_1 + \cdots + \alpha_{k-1} p_{k-1}$$

Nuevamente con un poco de álgebra llegaríamos a:

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

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

Como $\alpha_k = \dfrac{-r_k^T p_k}{p_k^T A p_k}$ y $\sigma_k = \dfrac{p_k^T A (x^* - x_0)}{p_k^T A p_k} = -\dfrac{p_k^T r_k}{p_k^T A p_k}$

Entonces $\alpha_k = \sigma_k$ y nos queda que efectivamente llegamos a x^* en a lo más n iteraciones.

Ahora, el método de gradiente conjugado es un método de direcciones conjugadas con la particularidad de que al ir generando los vectores conjugados, p_k se genera únicamente con p_{k-1}.

En particular, cada p_k se elige como la combinación lineal de $-r_k$ (residual) que es la máxima dirección de descenso y de p_{k-1}.

$$p_k = -r_k + \beta_k p_{k-1}$$

De esta manera, $\{p_1, p_2, \ldots, p_{n-1}\}$ son direcciones conjugadas y (por la pregunta anterior) linealmente independientes. En consecuencia, por el teorema expuesto y explicado anteriormente, el método de gradiente conjugado converge a la solución x^* en a lo más n pasos. $\perp$

2. BFGS

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

Solución:

P.d. $\underset{①}{H_{k+1}B_{k+1}} = \underset{②}{B_{k+1}H_{k+1}} = I$.

Tanto ① como ② se verifican de forma análoga, por lo que solo desarrollaremos ①.

Por definición tenemos:

① $B_{k+1}s_k = y_k$

② $B_{k+1} = (I-\rho_k y_k s_k^T)B_k(I-\rho_k s_k y_k^T)+\rho_k y_k y_k^T$

③ $\rho_k = \dfrac{1}{y_k^T s_k}$

④ $H_{k+1}y_k = s_k$

⑤ $H_{k+1} = H_k - \dfrac{H_k y_k y_k^T H_k}{y_k^T H_k y_k} + \dfrac{s_k s_k^T}{y_k^T s_k}$

Calculamos directamente:

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

$$= \underbrace{H_{k+1}(B_k-\rho_k y_k s_k^T B_k)(I-\rho_k s_k y_k^T)}_{A} + \underbrace{H_{k+1}\rho_k y_k y_k^T}_{B}$$

Desarrollamos primero A:

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

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

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

$H_kB_k = I$

(*) ρ_k es escalar ③

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

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

$$= \left[I - \frac{H_k y_k y_k^T}{y_k^T H_k y_k} \right] \left[I - \frac{s_k y_k^T}{y_k^T s_k} \right] = I - \cancel{\frac{H_k y_k y_k^T}{y_k^T H_k y_k}} - \frac{s_k y_k^T}{y_k^T s_k} + \cancel{\frac{H_k y_k (y_k^T s_k) y_k^T}{y_k^T H_k y_k (y_k^T s_k)}}$$

$$= I - \frac{s_k y_k^T}{y_k^T s_k} \quad (A)$$

Ahora desarrollamos B:

(4) $H_{k+1} y_k = s_k$

$$H_{k+1} \rho_k y_k y_k^T = \rho_k H_{k+1} y_k y_k^T = \frac{s_k y_k^T}{y_k^T s_k} \quad (B)$$

Nos queda entonces:

$$H_{k+1} B_{k+1} = I - \frac{s_k y_k^T}{y_k^T s_k} + \frac{s_k y_k^T}{y_k^T s_k} = I$$

$\therefore$ $H_{k+1} B_{k+1} = I$

$\therefore$ son inversas la una de la otra