Zara Ma. Ubaldo Galicia
166583

Final Análisis Ap.

① 1) $\langle p_0, p_1, \dots, p_{n-1} \rangle$ conjugados tq. $p_i^T A p_j = 0 \ \forall i \neq j$

Sup. que $p_j = \alpha_0 p_0 + \alpha_1 p_1 + \dots + \alpha_n p_n$

$p_i^T A p_j = 0 \Rightarrow \alpha_0 p_i^T A p_0 + \alpha_1 p_i^T A p_1 + \dots + \alpha_i p_i^T A p_i + \dots + \alpha_n p_i^T A p_n = 0$

$\Rightarrow \alpha_i p_i^T A p_i = 0 \Rightarrow \alpha_i = 0$ porque A es definida positiva

lo anterior ocurre $\forall i \neq j$

Por lo tanto, $\{p_0, \dots, p_{n-1}\}$ es linealmente independiente

2) En el método del gradiente conjugado, empezando en el punto x_0, generamos la sucesión $\{x_k\}$ de la forma:

$$x_{k+1} = x_k + \alpha_k p_k, \text{ donde } \alpha_k := -\frac{r_k^T p_k}{p_k^T A p_k}$$

Sabemos que encontrar la solución de $Ax=b$ es equivalente a minimizar $\phi(x) = \frac{1}{2} x^T A x - b^T x$, y sabemos que el minimizante unidimensional de $\phi(x)$ sobre $x_k + \alpha p_k$ es α_k,

Si después de n pasos, $x^* = \sigma_0 p_0 + \sigma_1 p_1 + \dots + \sigma_{n-1} p_{n-1}$ no es la solución que buscamos, hacemos otro paso del método

Sup $\hat{x} = x^* + \alpha_n p_n = x^* + \alpha_n(\delta_0 p_0 + \dots + \delta_{n-1} p_{n-1})$ porque $p_n \in \operatorname{span}\{p_0, \dots, p_{n-1}\}$

$= \lambda_0 p_0 + \dots + \lambda_{n-1} p_{n-1}$, pero una solución así la hubiéramos encontrado desde el paso anterior.

Así que, el método converge a la solución en a lo más n pasos

② 1) Segunda condición fuerte de Wolfe:

Sea $0 < c_1 < c_2 < 1$

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

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

$|a| \leq b$

$$s_k^T y_k = (x_{k+1} - x_k)^T (\nabla f_{k+1} - \nabla f_k) = \alpha_k p_k^T (\nabla f_{k+1} - \nabla f_k)$$

$$= \alpha_k p_k^T \nabla f_{k+1} - \alpha_k p_k^T \nabla f_k = \alpha_k (\nabla f_{k+1}^T p_k - \nabla f_k^T p_k)$$

$$|\nabla f(x_k + \alpha_k p_k)^T p_k| \leq c_2 |\nabla f_k^T p_k| \iff |\nabla f(\overbrace{x_k + \alpha_k p_k}^{x_{k+1}})^T p_k| \leq -c_2 \nabla f_k^T p_k$$

ya que $\nabla f_k^T p_k < 0$ porque p_k es dirección de descenso

Así, $\underbrace{c_2 \nabla f_k^T p_k \leq \nabla f(x_{k+1})^T p_k}_{(\star)} \leq -c_2 \nabla f_k^T p_k$

$$\boxed{\begin{aligned} &\cancel{\nabla f_{k+1} = y_k + \nabla f_k \Rightarrow \alpha_k \nabla f_k^T p_k = \alpha_k (y_k} \\ &\alpha_k \nabla f_{k+1}^T p_k = \alpha_k (y_k + \nabla f_k)^T p_k = y_k^T \alpha_k p_k + \alpha_k \nabla f_k^T p_k \\ &\qquad = y_k + \alpha_k \nabla f_k^T p_k \end{aligned}}$$

$$\cancel{0 \leq -c_2 \nabla f_k^T p_k - (-\nabla f_k + y_k)^T p_k = (1-c_2) \nabla f_k^T p_k - y_k^T p_k}$$

Por $(\star)$: $c_2 \alpha_k \nabla f_k^T p_k \leq \alpha_k (y_k + \nabla f_k)^T p_k$

$$\iff \quad 0 < y^T s_k + (1 - c_2) \alpha_k \nabla f_k^T p_k$$

$(1-c_2) > 0$ porque $c_2 < 1$ y $\nabla f_k^T p_k < 0$

> 0

$\therefore \; y_k^T s_k > 0$ ✗

$$0 < y_k^T p_k + \nabla f^T p_k - c_2 \nabla f^T p_k$$

$$0 < y_k^T p_k \; (1 - c_2) \nabla f^T p_k$$

2

$$(B_{k+1}H_{k+1})Y_k = B_{k+1}S_k = Y_k$$
$$(H_{k+1}B_{k+1})S_k = H_{k+1}Y_k = S_k$$

B_k simétrica y positiva definida

$$\left[(I-\rho_k Y_k S_k^T)B_k(I-\rho_k S_k Y_k^T)\right]^T$$
$$=(I-\rho_k S_k Y_k^T)^T B_k^T (I-\rho_k Y_k S_k^T)^T = (I-\rho_k Y_k S_k^T)B_k(I-\rho_k S_k Y_k^T)$$

$\therefore$ Es simétrica

$\rho_k = \frac{1}{Y_k^T S_k}$

$B_{k+1}S_k = Y_k$, $B_{k+1}^T = B_{k+1}$

$H_k = B_k^{-1}$

$H_{k+1}Y_k = S_k$

$Y^T S_k > 0$

$$\det\left[(I-\rho_k Y_k S_k^T)B_k(I-\rho_k S_k Y_k^T)\right] = \det(I-\rho_k Y_k S_k^T)\det(B_k)\det(I-\rho_k S_k Y_k^T)$$

$\det(B_k) \neq 0$ por ser definida positiva.

$$\det(I-\rho_k Y_k S_k^T) = \det\left[(I-\rho_k Y_k S_k^T)^T\right] = \det\left[I-\rho_k S_k Y_k^T\right]$$

No supe como demostrar que $\det(I-\rho_k S_k Y_k^T) \neq 0$

Si lo es, entonces $(I-\rho_k Y_k S_k^T)B_k(I-\rho_k S_k Y_k^T)$ es no singular

y $\rho_k Y_k Y_k^T$ es una modificacion de rango 1

donde $a = \sqrt{\rho_k}\,Y_k$ y $b^T = \sqrt{\rho_k}\,Y_k^T = a^T$

Así, por el lema de Sherman-Morrison-Woodbury tenemos que

$$B_{k+1}^{-1} = B_k^{-1} - \frac{B_k^{-1} \rho_k y_k y_k^T B_k^{-1}}{1 + \sqrt{\rho_k}\, y_k^T B_k^{-1} \sqrt{\rho_k}\, y_k} = H_k - \frac{\rho_k H_k y_k y_k^T H_k}{1 + \rho_k y_k^T H_k y_k}$$

$$= H_k - \frac{H_k y_k y_k^T H_k}{y_k^T s_k + y_k^T H_k y_k}$$