Examen Final José Alfredo Issa Barbará 173144

Supongamos que $p_0, p_1, \dots, p_\ell$ no son linealmente independientes.

Expresemos uno de ellos como combinación lineal del resto.

$$p_i = \alpha_0 p_0 + \cdots + \alpha_\ell p_\ell \quad \text{con} \quad \alpha_k \in \mathbb{R}, \quad k = 0, \dots, \ell.$$

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

Sea $k \in \{0, 1, \dots, \ell\}$

$$\Rightarrow \quad 0 = p_k^T A p_i = p_k^T A (\alpha_0 p_0 + \cdots + \alpha_\ell p_\ell)$$

$$= \alpha_0 p_k^T A p_0 + \cdots + \alpha_k p_k^T A p_k + \alpha_\ell p_k^T A p_\ell$$

$$= \alpha_k p_k^T A p_k$$ pues es el único caso donde $i = j$ y no se anula por hipótesis

Entonces tenemos que $\alpha_k p_k^T A p_k = 0$

Pero como A es positiva definida y los vectores son no nulos

$\Rightarrow \alpha_k = 0$

Y como esto lo hicimos para k arbitraria $\Rightarrow \alpha_k = 0 \quad \forall k = 0, 1, \dots, \ell$.

Y como $p_i = \alpha_0 p_0 + \alpha_1 p_1 + \cdots + \alpha_\ell p_\ell$, entonces

$$p_i = 0 p_0 + 0 p_1 + \cdots + 0 p_\ell = 0 \quad !$$

Pues p_i es un vector no nulo

$\therefore$ los vectores $p_0, p_1, \dots, p_\ell$ son linealmente independientes //.

Si la matriz A es diagonal,

las curvas de nivel de la función están alineadas con los ejes

y podemos encontrar el minimizador de la función al realizar

minimizaciones uni-dimensionales sobre las direcciones $e_1, e_2, \ldots, e_n$.

Cuando A no es diagonal,

las curvas de nivel de la función siguen siendo elipses pero ya no

están alineadas con los ejes.

En este caso, la solución no se encontraría en n pasos, incluso podría ser

que no se encuentre en un número finito de pasos.

Por lo que transformamos a A para que sea diagonal y ya

podemos volver a minimizar a través de las direcciones de las

coordenadas.

Cada minimización de coordenada determina correctamente uno de los

componentes de la solución que buscamos.

Es decir, después de n minimizaciones unidimensionales, se ha

realizado la minimización en span $\langle e_1, e_2, \ldots, e_n \rangle$, de donde

vemos que no puede ser en más de n iteraciones //.

Sabemos que

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

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

Entonces

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

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

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

$$= (I - \rho_k s_k y_k^T)(I - \rho_k y_k s_k^T) + B_k \rho_k s_k s_k^T$$

$$- \frac{B_k s_k s_k^T}{s_k^T B_k s_k} (I - \rho_k s_k y_k^T)(I - \rho_k y_k s_k^T) - \frac{B_k s_k s_k^T}{s_k^T B_k s_k} B_k \rho_k s_k s_k^T$$

$$+ y_k y_k^T \rho_k (I - \rho_k s_k y_k^T) H_k (I - \rho_k y_k s_k^T) + y_k y_k^T \rho_k \rho_k s_k s_k^T$$

$$= (I - \rho_k s_k y_k^T)(I - \rho_k y_k s_k^T) + B_k \rho_k s_k s_k^T$$

$$- (I - \rho_k s_k y_k^T)(I - \rho_k y_k s_k^T) - B_k \rho_k s_k s_k^T$$

$$+ \; y_k y_k^T \rho_k (I - \rho_k s_k y_k^T) H_k (I - \rho_k y_k s_k^T) + y_k y_k^T \rho_k \rho_k s_k s_k^T$$

$$= y_k y_k^T \rho_k (I - \rho_k s_k y_k^T) H_k (I - \rho_k y_k s_k^T) + y_k y_k^T \rho_k \rho_k s_k s_k^T$$

$$= H_k y_k y_k^T \rho_k \left[I - \rho_k y_k s_k^T - \rho_k s_k y_k^T + \rho_k s_k y_k^T \rho_k y_k s_k^T \right] + y_k y_k^T \rho_k \rho_k s_k s_k^T$$

$$= \frac{H_k y_k y_k^T}{y_k^T s_k} \left[I - \frac{y_k s_k^T}{y_k^T s_k} - \frac{s_k y_k^T}{y_k^T s_k} + \frac{s_k y_k^T}{y_k^T s_k} \frac{y_k s_k^T}{y_k^T s_k} \right] + \frac{y_k^T s_k}{y_k^T s_k} \cdot \frac{y_k s_k^T}{y_k^T s_k}.$$

$$= \frac{H_k y_k y_k^T}{y_k^T s_k} \left[I - 2I + I \right] + I$$

$$= \frac{H_k y_k y_k^T}{y_k^T s_k} \left[0 \right] + I$$

$$= I \quad //$$