3 valores por cada valor distinto de 0

val, i, j

/ /

166780

1. Gradiente conjugado

a) Demuestra que si $P_0, P_1, \dots, P_\ell$ satisfacen $P_i^T A P_j = 0 \ \forall i \neq j$, A simétrica y pos. def., ent. los vectores son LI.

Procedemos por contradicción

Supongamos que $P_0, P_1, \dots, P_\ell$ no son L.I, entonces podemos escribir a P_i como una combinación lineal de los otros vectores

$$P_i = \sigma_0 P_0 + \sigma_1 P_1 + \dots + \sigma_\ell P_\ell = \sum_{j=0}^{\ell} \sigma_j P_j \quad j \neq i$$

para algunos de los coeficientes σ_j.

Utilizando la propiedad $P_i^T A P_j = 0$ tenemos que:

$$0 = P_0^T A P_i = \sigma_0 P_0^T A P_0 + \dots + \sigma_\ell P_0^T A P_\ell$$
$$= \sigma_0 P_0^T A P_0 + 0 + \dots + 0$$
$$= \sigma_0 P_0^T A P_0$$

El mismo argumento es válido para los demás índices, lo que nos diría que la única explicación es que los coeficientes σ_k son cero y dado que $P_0, P_1, \dots, P_\ell$ no son L.I. $\Rightarrow$ los vectores son cero !

$\therefore$ $P_0, \dots, P_\ell$ son L.I

b) ¿Por qué el gradiente conjugado converge a lo más n iteraciones?

Si tenemos una matriz A de $n \times n$ con un conjunto de vectores $P_0, \dots, P_{n-1}$ que satisfacen la condición $P_i^T A P_j = 0$, por el inciso anterior, entonces son linealmente independientes y generan el espacio $\mathbb{R}^n$. Ahora, el algoritmo de gradiente conjugado es un método que nos genera direcciones conjugadas en una serie de pasos y al tener una matriz de $n \times n$, vamos a tener a lo más n direcciones y por tanto llegaremos al óptimo en a lo más n iteraciones.

Scribe

2. BFG

a) Demuestre que B_{k+1} y H_{k+1} son inversas una de la otra

$\boxed{\text{P.D.}}$ $B_{k+1} H_{k+1} = I$

Recordemos que:

1) $B_{k+1} s_k = y_k$

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

$$
\begin{aligned}
\Rightarrow B_{k+1} H_{k+1} &= B_{k+1}(I - \rho_k s_k y_k^T) H_k (I - \rho_k y_k s_k^T) + \rho_k s_k s_k^T \\
&= (B_{k+1} - \rho_k y_k y_k^T) H_k (I - \rho_k y_k s_k^T) + \rho_k y_k s_k^T \\
&= \left(B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k}\right) H_k (I - \rho_k y_k s_k^T) + \rho_k y_k s_k^T \\
&= \left(I - \frac{B_k s_k s_k^T}{s_k^T B_k s_k}\right)(I - \rho_k y_k s_k^T) + \rho_k y_k s_k^T \qquad (*) \\
&= I - \cancel{\frac{B_k s_k s_k^T}{s_k^T B_k s_k}} - \cancel{\rho_k y_k s_k^T} + \cancel{\frac{B_k s_k s_k^T}{s_k^T B_k s_k}} + \cancel{\rho_k y_k s_k^T} \\
&= I
\end{aligned}
$$

$(*)$ $\rho_k \dfrac{B_k s_k s_k^T}{s_k^T B_k s_k} y_k s_k^T = \dfrac{B_k s_k s_k^T}{s_k^T B_k s_k}$