Análisis Aplicado
Dr. Andreas Wachtel
Clase 16 (parte 1), 2 de Octubre de 2019

## Método del gradiente conjugado

La demostración del teorema que sigue usa Teorema 2 (clase 15) y las siguientes identidades:

$$\boldsymbol{r}_k = \boldsymbol{r}_{k-1} + \alpha_{k-1} A \boldsymbol{p}_{k-1} \tag{4}$$

$$\boldsymbol{p}_k \stackrel{\text{def}}{=\!=} -\boldsymbol{r}_k + \beta_k \boldsymbol{p}_{k-1} \tag{7e}$$

**Teorema 3** (Thm. 5.3, Nocedal)**.** *Si* $\boldsymbol{x}_k$ *no es la solución* $\boldsymbol{x}^\star$, *entonces*

$$\operatorname{span}\{\boldsymbol{r}_0, \boldsymbol{r}_1, \ldots, \boldsymbol{r}_k\} = \operatorname{span}\left\{\boldsymbol{r}_0, A\boldsymbol{r}_0, \ldots, A^k\boldsymbol{r}_0\right\}, \tag{8a}$$

$$\operatorname{span}\{\boldsymbol{p}_0, \boldsymbol{p}_1, \ldots, \boldsymbol{p}_k\} = \operatorname{span}\left\{\boldsymbol{r}_0, A\boldsymbol{r}_0, \ldots, A^k\boldsymbol{r}_0\right\}, \tag{8b}$$

$$\boldsymbol{p}_k^T A \boldsymbol{p}_i = 0 \qquad \textit{para} \qquad i = 0, \ldots, k-1. \tag{8c}$$

$$\boldsymbol{r}_k^T \boldsymbol{r}_i = 0 \qquad \textit{para} \qquad i = 0, \ldots, k-1, \tag{8d}$$

*Demostración.* Primero mostramos (8a) y (8b) por inducción.

El caso $k = 0$ es inmediato, puesto que $\boldsymbol{p}_0 = -\boldsymbol{r}_0$.

Entonces, como hipótesis de inducción tenemos que

$$\operatorname{span}\left\{\boldsymbol{r}_0, \boldsymbol{r}_1, \ldots, \boldsymbol{r}_{k-1}\right\} = \operatorname{span}\left\{\boldsymbol{r}_0, A\boldsymbol{r}_0, \ldots, A^{k-1}\boldsymbol{r}_0\right\}, \tag{$H_a$}$$

$$\operatorname{span}\left\{\boldsymbol{p}_0, \boldsymbol{p}_1, \ldots, \boldsymbol{p}_{k-1}\right\} = \operatorname{span}\left\{\boldsymbol{r}_0, A\boldsymbol{r}_0, \ldots, A^{k-1}\boldsymbol{r}_0\right\}. \tag{$H_b$}$$

Primero mostramos la relación $\subseteq$ de (8a) para $k$.

Hacemos el paso. Por las hipótesis $(H_a)$ y $(H_b)$ sabemos que

$$\boldsymbol{r}_{k-1}, \boldsymbol{p}_{k-1} \in \operatorname{span}\left\{\boldsymbol{r}_0, A\boldsymbol{r}_0, \ldots, A^{k-1}\boldsymbol{r}_0\right\}.$$

Entonces, $A\boldsymbol{p}_{k-1} = A\sum_{i=0}^{k-1} \gamma_i A^i \boldsymbol{r}_0 \in \operatorname{span}\left\{A\boldsymbol{r}_0, AA\boldsymbol{r}_0, \ldots, A^{k-1}A\boldsymbol{r}_0\right\}$ y (4) implican que

$$\boldsymbol{r}_k = \boldsymbol{r}_{k-1} + \alpha_{k-1} A\boldsymbol{p}_{k-1} \in \operatorname{span}\left\{\boldsymbol{r}_0, A\boldsymbol{r}_0, \ldots, A^k\boldsymbol{r}_0\right\}.$$

Con esto y la hipótesis $(H_a)$ concluimos que $\operatorname{span}\{\boldsymbol{r}_0, \boldsymbol{r}_1, \ldots, \boldsymbol{r}_k\} \subseteq \operatorname{span}\left\{\boldsymbol{r}_0, A\boldsymbol{r}_0, \ldots, A^k\boldsymbol{r}_0\right\}$.

Ahora, la relación $\supseteq$ de (8a). Por la hipótesis $(H_b)$ y identidad (4) sabemos que

$$A^k\boldsymbol{r}_0 = A(A^{k-1})\boldsymbol{r}_0 = A\sum_{j=0}^{k-1}\gamma_i \boldsymbol{p}_i = \sum_{j=0}^{k-1}\gamma_i A\boldsymbol{p}_i = \sum_{j=0}^{k-1}\gamma_i \frac{1}{\alpha_i}\left(\boldsymbol{r}_{i+1} - \boldsymbol{r}_i\right)$$

Entonces, $A^k\boldsymbol{r}_0 \in \operatorname{span}\{\boldsymbol{r}_0, \boldsymbol{r}_1, \ldots, \boldsymbol{r}_k\}$ y con la hipótesis $(H_a)$ concluimos

$$\operatorname{span}\{\boldsymbol{r}_0, \boldsymbol{r}_1, \ldots, \boldsymbol{r}_k\} \supseteq \operatorname{span}\left\{\boldsymbol{r}_0, A\boldsymbol{r}_0, \ldots, A^k\boldsymbol{r}_0\right\}.$$

Hemos mostrado que (8a) también se cumple para $k$.

Ahora, mostramos (8b). Empezamos utilizando (7e), ...

$$\begin{aligned}
\operatorname{span}\{\boldsymbol{p}_0, \boldsymbol{p}_1, \ldots, \boldsymbol{p}_k\} &= \operatorname{span}\{\boldsymbol{p}_0, \boldsymbol{p}_1, \ldots, \boldsymbol{p}_{k-1}, \boldsymbol{r}_k\} && \text{, por (7e),}\\
&= \operatorname{span}\left\{\boldsymbol{r}_0, A\boldsymbol{r}_0, \ldots, A^{k-1}\boldsymbol{r}_0, \boldsymbol{r}_k\right\} && \text{, por } (H_b),\\
&= \operatorname{span}\{\boldsymbol{r}_0, \boldsymbol{r}_1, \ldots, \boldsymbol{r}_{k-1}, \boldsymbol{r}_k\} && \text{, por } (H_a),\\
&= \operatorname{span}\left\{\boldsymbol{r}_0, A\boldsymbol{r}_0, \ldots, A^{k-1}\boldsymbol{r}_0, A^k\boldsymbol{r}_0\right\} && \text{, por (8a) para } k.
\end{aligned}$$

□

*Demostración.* Ahora mostramos (8c), *i.e.*, que los $\boldsymbol{p}_i$ son $A$ -conjugados por inducción. Para $k = 1$ notamos que $\boldsymbol{p}_1$ fue construido tal que $\boldsymbol{p}_1^T A\boldsymbol{p}_0 = 0$. Por lo cual, la hipótesis de inducción es que (8c) es valida para $k-1$. Con el paso de inducción llegamos a $k$ como sigue:

Multiplicamos (7e) por $A\boldsymbol{p}_i$, $(i = 0, \ldots, k-1)$, es decir

$$\begin{aligned}
\boldsymbol{p}_k^T A\boldsymbol{p}_i &= \big(-\boldsymbol{r}_k + \beta_k \boldsymbol{p}_{k-1}\big)^T A\boldsymbol{p}_i\\
&= -\boldsymbol{r}_k^T A\boldsymbol{p}_i + \beta_k \boldsymbol{p}_{k-1}^T A\boldsymbol{p}_i = \begin{cases} 0 & \text{, para } i = k-1 \text{ por construcción,}\\ -\boldsymbol{r}_k^T A\boldsymbol{p}_i & \text{, para } i \le k-2 \text{ por la hipótesis de Ind,}\end{cases}
\end{aligned}$$

puesto que para $i \le k-2$, la simetría de $A$ y la hipótesis ((8c) para $k-1$) implican que $\{\boldsymbol{p}_0, \ldots, \boldsymbol{p}_{k-1}\}$ son $A$ -conjugados. Con esto y Teorema 2 concluimos que

$$\boldsymbol{r}_k^T \boldsymbol{p}_i = 0 \qquad \text{para} \qquad i = 0, \ldots, k-1 \qquad \text{(Es la ecuación (5) de nuevo).}$$

Luego (8b) implica que

$$A\boldsymbol{p}_i = A\sum_{j=0}^{i} A^j \boldsymbol{r}_0 = \sum_{j=0}^{i} A^{j+1}\boldsymbol{r}_0 \in \operatorname{span}\{\boldsymbol{p}_0, \boldsymbol{p}_1, \ldots, \boldsymbol{p}_{i+1}\},$$

por lo cual, concluimos que $\boldsymbol{r}_k^T A\boldsymbol{p}_i = 0$ para $i \le k-2$. Con esto (8c) queda demostrado. □

*Demostración.* Para mostrar (8d), usamos que por (8c) el conjunto $\{\boldsymbol{p}_0, \ldots, \boldsymbol{p}_k\}$ es $A$ -conjugado. Con Teorema 2 deducimos que

$$\boldsymbol{r}_k^T \boldsymbol{p}_i = 0 \qquad \text{para} \qquad i < k \qquad \text{y para} \qquad k = 1, 2, \ldots, n-1.$$

Usando (7e) deducimos que $\boldsymbol{r}_i \in \operatorname{span}\{\boldsymbol{p}_i, \boldsymbol{p}_{i-1}, \boldsymbol{p}_0\}$, por lo cual, esa ortogonalidad muestra

$$\boldsymbol{r}_k^T \boldsymbol{r}_i = 0 \qquad \text{para} \qquad i = 0, \ldots, k-1,$$

es decir, hemos mostrado (8d). □