Cuadrados Mínimos

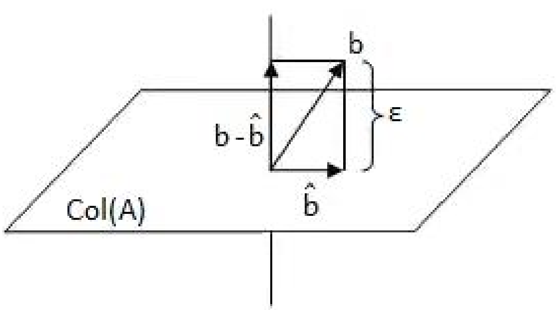
by gira

09/10/2009

Sean $A \in R^{mxn}$, $b \in R^m$ y $x \in R^n$:

el sistema $Ax = b$ entonces es compatible $\iff$ $b \in Col(A)$

Ahora, sea el sistema $Ax = b$ tal que $b \notin Col(A)$ (uno incompatible) entonces queremos la solución de este sistema (un $\widehat{x}$) de forma tal de encontrar un $\widehat{b}$ que perteneza al col(A) y sea lo más "parecido" a b posible, o mejor dicho, de forma que el vector $\widehat{b}$ sea el que más se aproxime a b (y el "error" sea el mínimo), es decir de forma que $\widehat{b} = p_{Col(A)}(b)$



Ejemplo en R3

El error ε es igual a $\left\|b - \widehat{b}\right\|$.

Entonces el objetivo es encontrar la solución a este sistema: $A\widehat{x} = \widehat{b} = p_{Col(A)}(b)$. La solución $\widehat{x}$ recibe el nombre de *"solución por cuadrados mínimos"* del sistema $Ax = b$.

Entonces veamos ahora como encontrar $\widehat{x}$ sin necesidad de calcular la proyección de b sobre $Col(A)$:

$$A\widehat{x} = p_{Col(A)}(b) \iff \begin{cases} A\widehat{x} \in Col(A) \\ b - A\widehat{x} \perp Col(A) \end{cases}$$

Recordemos que para matrices reales y con el prod. interno canónico se cumple que:

$[Col(A)]^{\perp} = Nul\left(A^T\right)$

$\implies b - A\widehat{x} \in Nul\left(A^T\right) \implies A^T(b - A\widehat{x}) = 0 \implies A^T b - A^T A\widehat{x} = 0 \iff$

$$\boxed{A^T A\widehat{x} = A^T b} \tag{1}$$

A este sistema se lo llama *"Ecuaciones normales de cuadrados mínimos"*. La solución de este sistema es la solución por cuadrados mínimos que buscabamos.

Sin embargo todo no termina aquí. Para este sistema tenemos dos casos: que sea compatible determinado **(SCD)** o compatible indeterminado **(SCI)**, y cada caso tiene sus propiedades; veamos:

Caso 1: Si $A^T A$ es inversible, entonces no está nada mal que cambiemos el sistema a esta forma: $\widehat{x} = \left(A^T A\right)^{-1} A^T b$, De esta forma podemos afirmar que el sistema es compatible determinado y la solución es única. La matriz $A^{\#} = \left(A^T A\right)^{-1} A^T$ es llamada "matriz pseudoinversa de A".

Ahora, observen lo siguiente: Si $A^T A$ es inversible entonces como recordarán $rg(A^T A) = n$, pero además existe una propiedad que dice que $rg(A^T A) = rg(A)$ (ver ej. 21 de la guía), entonces este caso podría resumirse de la siguiente forma:

Ecs. normales son SCD $\iff$ $\widehat{x}$ es único $\iff$ $A^T A$ es inversible $\iff$ $rg(A^T A) = rg(A) = n$ $\iff$ Las columnas de A son LI (o $det(A) \neq 0$, etc) $\iff$ $\widehat{x} = \left(A^T A\right)^{-1} A^T b$ $\iff$ $\widehat{x} = A^{\#} b$

Propiedades de la Pseudoinversa:

1. $A^{\#} \in R^{nxm}$

2. $A^{\#} A = I$ (pues $A^{\#} A = \left(A^T A\right)^{-1} A^T A = I$)

3. $AA^{\#} = P_{Col(A)}(b)$ (pues $\widehat{x} = A^{\#} b \rightarrow A\widehat{x} = AA^{\#} b \rightarrow p_{Col(A)}(b) = AA^{\#} b \rightarrow P_{Col(A)}(b) = AA^{\#}$)

Caso 2: Si $A^T A$ no es inversible, entonces $\widehat{x}$ no es único y las ecs. normales son un SCI. Además las columnas de A son LD y su determinante es 0 (pues $rg(A) = rg(A^T A)$). Sin embargo podemos encontrar una expresión de $\widehat{x}$ de la siguiente forma:

$\widehat{x} = \widehat{x}_p + \widehat{x}_h$

La solución $\widehat{x}$ la podemos expresar como la suma de una particular ($\widehat{x}_p$) y una homogénea ($\widehat{x}_h$). La homogénea sería la solución de $A^T A\widehat{x} = 0$, y la particular es simplemente una solución de las infinitas que hay. Podemos observar entonces que $\widehat{x}_h \in Nul(A^T A)$, pero (ver ej. 21 de la guía) $Nul(A^T A) = Nul(A)$, entonces $\widehat{x}_h \in Nul(A)$!.

Algunas propiedades a tener en cuenta:

1. Sea $\widehat{x}$ solución por cuadrados mínimos del sistema $Ax = b \implies \|A\widehat{x}\| \leq \|b\|$

2. $\widehat{x} = 0$ es solución por cuadrados mínimos del sistema $Ax = b \iff b \perp Col(A)$