Moore-Penrose Pseudoinverse: Classical least squares problem

J. Antonio Ortega UNAM

Lemma: Let x be a vector and $\mathcal{L}$ is a linear manifold in $\mathbb{R}^n$ (*i.e.,* if x, $y \in \mathcal{L}$, then $\alpha x + \beta y \in \mathcal{L}$ for any scalars α, β). Then if

$$x = \hat{x} + \tilde{x}$$

where $\hat{x} \in \mathcal{L}$ and $\tilde{x} \perp \mathcal{L}$, then $\tilde{x}$ is *"nearest"* to x, or, in other words, it is the projection of x to the manifold $\mathcal{L}$.

Proof.: For any $y \in \mathcal{L}$ we have

$$
\begin{aligned}
\|x - y\|^2 &= \|\hat{x} + \tilde{x} + y\|^2 = \|(\hat{x} - y) + \tilde{x}\|^2 \\
&= \|(\hat{x} - y)\|^2 + 2(\hat{x} - y,\, \tilde{x}) + \|\tilde{x}\|^2 = \|(\hat{x} - y)\|^2 + \|\tilde{x}\|^2 \\
&\geq \|\tilde{x}\|^2 = \|x - \hat{x}\|^2
\end{aligned}
$$

Theorem: Let z be an $n-$dimensional real vector and $H \in \mathbb{R}^{n \times m}$.

- 1. There is always a vector, in fact a unique vector $\hat{x}$ of minimal (Euclidian norm), which minimizes
$$\|z - Hx\|^2.$$

- 2. The vector $\hat{x}$ is the unique vector in the range
$$\mathcal{R}(H^T) := \{x : x = H^T z, z \in \mathbb{R}^n\}$$
which satisfies the equation
$$H\hat{x} = \hat{z}$$
where $\hat{z}$ is the projection of z on $\mathcal{R}(H)$.

Proof.: We can write

$$z = \hat{z} + \tilde{z}$$

where $\hat{z}$ is the projection of z on the kernel (*null space*)

$$\mathcal{N}(H^T) := \{z \in \mathbb{R}^n | H^T z = 0\}.$$

Since $Hx \in \mathcal{R}(H)$ for any $x \in \mathbb{R}^m$, it follows that

$$\hat{z} - Hx \in \mathcal{R}(H)$$

and, since $\tilde{z} \in \mathcal{R}^{\perp}(H)$,

$$\tilde{z} \perp \hat{z} - Hx, \therefore$$

$$\begin{aligned} \|z - Hx\|^2 &= \|(\hat{z} - Hx) + \tilde{z}\|^2 \\ &= \|\hat{z} - Hx\|^2 + \|\tilde{z}\|^2 \geq \|\tilde{z}\|^2 = \|z - \hat{z}\|^2 \end{aligned}$$

$\|z - Hx\|^2 \geq \|z - \hat{z}\|^2$ This low bound is attainable since $\hat{z}$, being the range of H, is the afterimage of some x^*, that is, $\hat{z} = Hx^*$.

1. Let us show that x^* has a minimal norm.

$$x^* = \hat{x}^* + \tilde{x}^*$$

where $\hat{x}^* \in \mathcal{R}(H^\perp)$ and $\tilde{x}^* \in \mathcal{N}(H)$.

Thus, $Hx^* = H\hat{x}$ we have $\|z - Hx^*\|^2 = \|z - H\hat{x}\|^2$ and $\|x^*\|^2 \geq \|\hat{x}^*\|^2$.

So, x^* may be selected equal to $\hat{x}^*$.

2. Show that $x^* = \hat{x}^*$ is unique. Suppose that $Hx^* = Hx^{**} = \hat{z}$. Then

$$(x^* - x^{**}) \in \mathcal{R}(H)$$

But, $H(x^* - x^{**}) = 0$, therefore $(x^* - x^{**}) \in \mathcal{N}(H) = \mathcal{R}^{\perp}(H^{\perp})$
Thus, $(x^* - x^{**})$ is orthogonal to itself! i.e., $\|x^* - x^{**}\|^2 = 0 \implies x^* = x^{**}$.

Corollary: $\|z - Hx\|^2$ is minimized by x_0 if and only if $Hx_0 = \hat{z}$ where $\hat{z}$ is the projection of z on $\mathcal{R}(H)$.

Corollary: There is always an $n-$dimensional vector y such that

$$\|z - HH^T y\|^2 = \inf_x \|z - Hx\|^2,$$

and if

$$\|z - Hx_0\|^2 = \inf_x \|z - Hx\|^2,$$

then

$$\|x_0\|^2 \geq \|H^T y\|^2$$

with strict inequality unless $x_0 = H^T y$. The vector y satisfies the equation $HH^T y = \hat{z}$

Theorem: Among those vectors x, which minimize $\|z - Hx\|^2$, $\hat{x}$, the one having minimal norm, is the unique vector of the form:

$$\hat{x} = H^T y,$$

satisfying

$$H^T H\hat{x} = H^T z.$$