Derivation of Unconstrained Closed-Form Least Squares Solution

Greg Reynolds

28th March 2007

It is desired to solve the following problem:

$$\min_{\boldsymbol{x}} \frac{1}{2} \parallel \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b} \parallel_2^2 \tag{1}$$

where $\boldsymbol{A} \in \mathbb{C}^{m \times n}, \boldsymbol{x} \in \mathbb{C}^n$ and $\boldsymbol{b} \in \mathbb{C}^m$. The $\frac{1}{2}$ in (1) is a convenience factor. To obtain the solution, $\boldsymbol{x}$, define the scalar expression for error:

$$\begin{aligned} e &= \frac{1}{2} \parallel \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b} \parallel_2^2 \\ &= \frac{1}{2}(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b})^H(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}) \\ &= \frac{1}{2}\left[(\boldsymbol{A}\boldsymbol{x})^H\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}^H\boldsymbol{A}\boldsymbol{x} - (\boldsymbol{A}\boldsymbol{x})^H\boldsymbol{b} + \boldsymbol{b}^H\boldsymbol{b}\right] \\ &= \frac{1}{2}\left[\boldsymbol{x}^H\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}^H\boldsymbol{A}\boldsymbol{x} - \boldsymbol{x}^H\boldsymbol{A}^H\boldsymbol{b} + \boldsymbol{b}^H\boldsymbol{b}\right] \end{aligned} \tag{2}$$

then take the derivative. Since $\boldsymbol{x}$ has n-elements, the derivative will be n-dimensional. Standard results of derivatives of vectors expressions [?] can be applied to each part of (2):

$$\frac{\partial}{\partial \boldsymbol{x}}\boldsymbol{x}^H\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{x} = 2\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{x} \tag{3}$$

$$\frac{\partial}{\partial \boldsymbol{x}}\boldsymbol{b}^H\boldsymbol{A}\boldsymbol{x} = (\boldsymbol{b}^H\boldsymbol{A})^H = \boldsymbol{A}^H\boldsymbol{b} \tag{4}$$

$$\frac{\partial}{\partial \boldsymbol{x}}\boldsymbol{x}^H\boldsymbol{A}^H\boldsymbol{b} = \boldsymbol{A}^H\boldsymbol{b} \tag{5}$$

$$\frac{\partial}{\partial \boldsymbol{x}}\boldsymbol{b}^H\boldsymbol{b} = 0 \tag{6}$$

which enables the writing of the derivative of the error with respect to $\boldsymbol{x}$ as:

$$\begin{aligned} \frac{\partial e}{\partial \boldsymbol{x}} &= \frac{1}{2}\left[2\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{x} - 2\boldsymbol{A}^H\boldsymbol{b}\right] \\ &= \boldsymbol{A}^H\left[\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\right] \\ &= \boldsymbol{A}^H\boldsymbol{A}\boldsymbol{x} - \boldsymbol{A}^H\boldsymbol{b} \end{aligned} \tag{7}$$

The minimum e occurs when its derivative is zero, so:

$$\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{x} - \boldsymbol{A}^H\boldsymbol{b} = 0 \tag{8}$$

which can be rearranged first to:

$$\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^H\boldsymbol{b} \tag{9}$$

and finally to:

$$\boldsymbol{x} = (\boldsymbol{A}^H\boldsymbol{A})^{-1}\boldsymbol{A}^H\boldsymbol{b} \tag{10}$$

The expression $(\boldsymbol{A}^H\boldsymbol{A})^{-1}\boldsymbol{A}^H$ is often called the *Moore-Penrose* inverse of $\boldsymbol{A}$.