14 Lecture 14: The least squares problems (27/03/2020)

14.1 Least square solution

The least square solution is a method of solving inconsistent linear systems.

Theorem 1 *$A\vec{x} = \vec{b}$ is consistent iff $\vec{b} \in Col(A)$, since $\vec{b}$ can then be written as a linear combination of the columns of A.*

Thus a system is inconsistent if $\vec{b} \notin \mathrm{Col}(A)$. To solve an inconsistent linear system for the closest possible solution we solve for the vector $\hat{b}$ which is the projection of the vector $\vec{b}$ onto the column space of A[1]. The system then becomes:

$$A\hat{x} = \hat{b} \tag{1}$$

Note that the vector $\vec{x}$ is a different vector then the vector $\hat{x}$. $\vec{x}$ would be the actual solution to the system which we cannot find and $\hat{x}$ is the closest approximate solution to the system. Since the vector $\hat{x}$ is defined to be part of the column space of A we know for a fact that this system is consistent. An

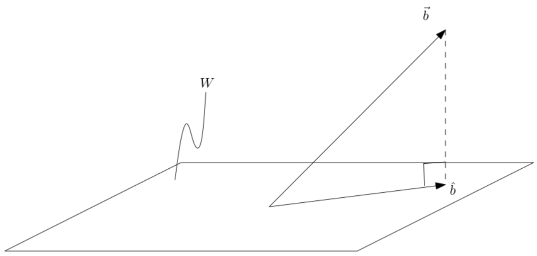

Figure 1: Graphical representation of the projection of the vector $\vec{x}$ onto the column space of A ($W = \mathrm{Col}(A)$)

interesting thing to note is the following theorem, which offers a different way of computing the least squares solution.

Theorem 2 *The following statements are equivelant:*

- *$\hat{x}$ is the LSQ solution to $A\vec{x} = \vec{b}$.*
- *$\hat{x}$ is the (actual) solution to the equation $A^TA\hat{x} = A^T\vec{b}$.*

14.2 LSQ Error

As mentioned before the solution to the system $A\hat{x} = \hat{b}$ is not an exact solution, but rather the closest possible solution. Not all approximate solutions are made equal. Some will have a large error while some have a smaller error. We can find the error by checking the distance from the vector $\vec{b}$ to the new projected vector $\hat{b}$. This distance represents how much the system had to be adjusted to make it consistent. When the distance is very large a big adjustment to the system had to be made to force consistency. If the distance is very small only a small adjustment had to be made to make the system consistent.

$$\mathrm{dist}(\vec{b}, \hat{b}) = ||\vec{b} - \hat{b}|| = ||\vec{b} - A\hat{x}|| \tag{2}$$

[1] $\mathrm{proj}_{\mathrm{Col}(A)}(\vec{b})$

14.3 Example with linear regression

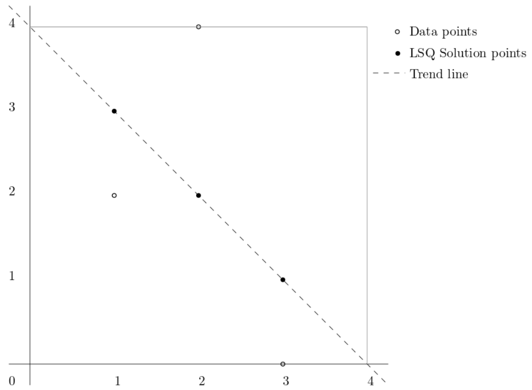

Figure 2: The graph of the inconsistent linear system and the closest consisten solution

$$\text{let:}A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}, \ \vec{x} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}, \ \vec{b} = \begin{bmatrix} 2 \\ 4 \\ 0 \end{bmatrix} \tag{1}$$

The column space of A will then be:

$$\text{Col}(A) = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right\} \tag{2}$$

We now apply the Gram-Schmidt process to find an orthogonal basis for the column space of A:

$$\vec{v}_1 = \vec{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \tag{3}$$

$$\vec{v}_2 = \vec{x}_2 - \frac{\langle \vec{x}_2 | \vec{v}_1 \rangle}{\langle \vec{v}_1 | \vec{v}_1 \rangle} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \frac{6}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \tag{4}$$

The orthogonal basis for will then be $\{\vec{v}_1, \vec{v}_2\}$ which is:

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\} \tag{5}$$

The projection of vector b onto the column space of A $\hat{b}$ will then be:

$$\begin{aligned} \text{proj}_{\text{Col}(A)}(\vec{b}) = \hat{b} &= \frac{\langle \vec{b} | \vec{v}_1 \rangle}{\langle \vec{v}_1 | \vec{v}_1 \rangle} \vec{v}_1 + \frac{\langle \vec{b} | \vec{v}_2 \rangle}{\langle \vec{v}_2 | \vec{v}_2 \rangle} \vec{v}_2 \\ \hat{b} &= \frac{6}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \frac{-2}{2} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \\ \hat{b} &= \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \end{aligned} \tag{6}$$

The LSQ solution to the problem will then become:

$$A\hat{x} = \hat{b} \Rightarrow \left[\begin{array}{cc|c} 1 & 1 & 3 \\ 1 & 2 & 2 \\ 1 & 3 & 1 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 1 & 3 \\ 0 & 1 & -1 \\ 0 & 2 & -2 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 0 & 4 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{array}\right] \tag{7}$$

$$\text{thus: } \hat{\beta}_0 = 4,\ \hat{\beta}_1 = -1 \Rightarrow \hat{x} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \end{bmatrix} \tag{8}$$

Where equation (8) is the closest consistent solution for the linear system.