6.5. Least Square problems

Def: If $A$ is an $m \times n$ matrix and $b$ is in $\mathbb{R}^m$, a least-square solution of 

$$Ax = b$$

is a vector $\hat{x}$ in $\mathbb{R}^n$ such that 

$$\|b - A \hat{x}\| \leq \|b - Ax\| \quad \text{for all } x \text{ in } \mathbb{R}^n$$

Remark: If $b$ is in col $A$, then $\hat{x}$ is any solution of $Ax = b$ since then 

$$\|b - A \hat{x}\| = 0 \leq \|b - Ax\| \quad \text{for all } x \text{ in } \mathbb{R}^n$$

If $b$ is not in col $A$, then $\hat{x}$ is any solution of $Ax = \text{proj}_A b$ since in this case 

$$\|b - A \hat{x}\| = \|b - \text{proj}_A b\| \leq \|b - Ax\|$$

Note that if $b$ is in col $A$ then $\text{proj}_A b = b$
Conclusion: All least square solutions of $Ax = b$ are given by the solutions of the consistent equation:

(1)  $Ax = \text{proj}_{\text{col } A} \ b$

Remark: To solve (1) one needs to compute $\text{proj}_{\text{col } A} \ b$. This requires an orthogonal basis in $\text{col } A$ which can be obtained by orthogonalizing the basis of $\text{col } A$ from by the linear independent columns of $A$, see section 6.4 (not required).

The following theorem provides a way around computing $\text{proj}_{\text{col } A} \ b$:

Theorem: All least square solutions of $Ax = b$ are given by the solutions of the consistent equation:

(2)  $A^TAx = A^Tb$

Proof. First we show that any solution of (1) is a solution of (2):

If $x$ solves (1) then

$b - Ax = 0 - \text{proj}_{\text{col } A} b = L \cdot a_j$ where

$a_j$ is any column of $A$. 

Hence \( a^T (b - A\hat{x}) = 0 \)

\[\Rightarrow A^T (y - A\hat{x}) = 0 \text{ since } A^T \text{ is formed by the columns of } A\]

\[\Rightarrow A^T A\hat{x} = A^T y \text{ so } \hat{x} \text{ solves (2)}\]

What we show is that any solution of (2) is a solution of (1):

Let \( \hat{x} \) be a solution of (2) then

\[A^T (y - A\hat{x}) = 0\]

So \( b - A\hat{x} \) is orthogonal to each column of \( A \).

Since \( \text{col } A = \text{span } \{ \text{columns of } A \} \) we deduce that \( b - A\hat{x} \) is orthogonal to any vector in \( \text{col } A \).

By definition of \( P_{\text{col } A} b \) we get

\[A\hat{x} = P_{\text{col } A} b \text{ so } \hat{x} \text{ satisfies (1)}\]

In conclusion, the set of solutions of (2) is the same as the set of solutions of (1) which is the set of all least squares solutions of \( A\hat{x} = b \).
Example  Find a least square solution of

\[
\begin{bmatrix}
2 & 1 \\
-2 & 0 \\
2 & 3
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2
\end{bmatrix} =
\begin{bmatrix}
-5 \\
8 \\
1
\end{bmatrix}
\]

\[
E_2 (2) \text{ becomes: }
\begin{bmatrix}
2 & -2 & 2 \\
1 & 0 & 3 \\
1 & 0 & 3
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2
\end{bmatrix} =
\begin{bmatrix}
2 & -2 & 2 \\
1 & 0 & 3 \\
1 & 0 & 3
\end{bmatrix}
\begin{bmatrix}
-5 \\
8 \\
1
\end{bmatrix}
\]

\[
\begin{bmatrix}
12 & 8 \\
8 & 10
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2
\end{bmatrix} =
\begin{bmatrix}
-24 \\
-2
\end{bmatrix}
\]

\[
\begin{bmatrix}
x_1 \\
x_2
\end{bmatrix} =
\begin{bmatrix}
12 & 8 \\
8 & 10
\end{bmatrix}^{-1}
\begin{bmatrix}
-24 \\
-2
\end{bmatrix} = \frac{1}{56}
\begin{bmatrix}
10 & -8 \\
-8 & 12
\end{bmatrix}
\begin{bmatrix}
-24 \\
-2
\end{bmatrix}
\]

\[
\begin{bmatrix}
-4 \\
3
\end{bmatrix}
\]

Remark  Solutions of (1) (or (2)) are not always unique. There is a unique least square solution if and only if $A$ has linearly independent columns, or equivalently, $A^T A$ is invertible. In
all other sons are usually chooses the solution with the smallest length.