Numerical Linear Algebra
Unit 6: Least Squares Problems

Numerical Analysis, Lund University

Claus Führer and Philipp Birken
Overdetermined System

Let $A \in \mathbb{R}^{m \times n}$ with $m > n$.

\[ Ax = b \]

has a solution iff $b \in \text{range}(A)$.

Otherwise

\[ 0 \neq r = b - Ax \]

$r$ called the residual.
$x^*$ is called a generalized solution of $Ax = b$ (or its solution in the least squares sense) if

$$x^* = \arg\min_{x \in \mathbb{R}^n} \| b - Ax \|_2$$

Solution: Let $P_A$ be the orthogonal projection from $\mathbb{R}^m$ onto range$(A)$.

Then, $x^*$ is a solution of $Ax = P_A b$.

Needs not to be unique.
Example: Polynomial Fitting

Let us be given $m$ pairs of data $(x_i, y_i)$.

Find among all polynomials

$$p(x) = c_{n-1}x^{n-1} + \cdots + c_1x + c_0$$

of max-degree $n - 1$ the polynomial $p^*$ which best “describes” the data:

$$p^* = \arg\min_{p \in \mathcal{P}^{n-1}} \sum_{i=1}^{m} |p(x_i) - y_i|^2$$
Vandermonde System

\[
\begin{pmatrix}
1 & x_1 & \cdots & x_1^{n-1} \\
1 & x_2 & \cdots & x_1^{n-1} \\
1 & \vdots & & \vdots \\
1 & x_{m-1} & \cdots & x_{m-1}^{n-1} \\
1 & x_m & \cdots & x_m^{n-1}
\end{pmatrix}
\begin{pmatrix}
c_0 \\
c_1 \\
c_2 \\
c_{n-1}
\end{pmatrix}
\begin{pmatrix}
y_1 \\
y_2 \\
\vdots \\
y_{m-1} \\
y_m
\end{pmatrix}
\]

Check the use of \texttt{polyfit} in MATLAB or Python. See even the command \texttt{vander}.
Solution of the Vandermonde System

If $x_i \neq x_j$ for $i \neq j$ the Vandermonde matrix $V$ has full rank. An orthogonal projector on range $V$:

$$P_V = V(V^TV)^{-1}V^T$$

Thus,

$$b^* = P_v b = V(V^TV)^{-1}V^Tb \implies Vx^* = V(V^TV)^{-1}V^Tb$$

From this we get the solution

$$V^TVx^* = V^TV(V^TV)^{-1}V^Tb = V^Tb$$

$$\implies x^* = (V^TV)^{-1}V^Tb$$
Normal equations

The equation

$$A^T A x^* = A^T b$$

is called the *normal equation* of the least squares problem

$$x^* = \text{argmin}_{x \in \mathbb{R}^n} \| b - Ax \|_2$$
Example for polynomial fit: Oat Porridge Problem
We pose the following questions:

- How much water is needed for three portions?
- How much water is needed for 300 portions?

To answer these questions, we first have to set up a correct mathematical model and then make a least squares fit to determine the unknown parameters in the model.

The discussion of the different models will take place in the lecture - not here.
Definition
If $A \in \mathbb{R}^{m \times n}$ with $m \geq n$ has full rank, the matrix

$$A^\dagger := (A^T A)^{-1} A^T$$

is called the *pseudoinverse* of $A$.

The least-squares solution of $Ax = b$ is $x = A^\dagger b$.

It is the classical solution of $Ax = P_A b$ with the projection matrix $P_A := AA^\dagger$. 

Pseudoinverse