ECS230

Summary 5

Least Squares Problem

Least squares problem: $X \in \mathbb{R}^{m \times n}$ $y \in \mathbb{R}^m$ $m > n$
The problem $Xb = y$ does not in general have a solution (more equations than unknowns).
The least squares (LS) problem consists of finding a vector $b \in \mathbb{R}^n$ such that $||Xb - y||$ is minimized.
Data-fitting interpretation: Find the best description of the vector y as a linear combination of the columns of X.

Normal equations:
Assuming $\text{rank}(X) = n$, the residual norm $||Xb - y||$ is minimized if b is a solution of the Normal Equations

$$X^T X b = X^T y$$

Solution of the Normal Equations:
i) Compute $A = X^T X$ (A is symmetric positive definite)
ii) Cholesky factorization: $A = LL^T$
iii) Define $u = L^T b$. Solve $Lu = X^T y$. Solve $L^T b = u$

Polynomial fit:
Given a set of pairs (x_i, y_i) $i = 1, \ldots, p$, find a polynomial of degree $d < p$ that approximates the data in the least-squares sense.
Solve the Normal Equations with $X \in \mathbb{R}^{p \times (d+1)}$, and $x_{ij} = (x_i)^{(j-1)}$.

Accuracy:
Condition number $\kappa(X^T X) = \kappa(X)^2$, where

$$\kappa(X) = ||X^+||\,||X||$$

and X^+ is the Moore-Penrose pseudoinverse of X. The condition number of $X^T X$ can be large. The QR factorization can be used to solve the LS problem without having to form the product $X^T X$.