Math 307

Lecture 14: Least Squares Approximation

2025-10-21

0.1 Least Squares Approximation

- QR decomposition, meant to find an orthonormal basis for the column space of the column space of A.
- Consider A is n x m matrix and rank$(A) =$ m (and $m \geq n$) and let $b \in \mathbb{R}$ This is a wide matrix. Then either $Ax = b$ has a unique solution or no solution. What if there is no solution? We need to find x such that Ax is a close to b as possible. Why we do Least Square Approximation, (No solution).
- *Definition* : The least squares approximation of $Ax \neq b$ is $x \in \mathbb{R}$ which minimizes $\|Ax - b\|$. We can think of this as a projection, we are not projecting x
- Note: There are many ways to compute the least squares approximation: Normal equations, $A^T Ax = A^T b$, QR decomposition $R_1 x = Q^T b$,SVD $x = A^T b$, or to minimize we can use calculus too, but we care most about the first two for now.
- *Theroem* (Normal equations) : The least squares approximation of $Ax \risingdotseq b$ is the solution of $A^T Ax = A^T b$, reminds us of the projection theorem

  Proof (Recall projection theorem)

  - So we know that $Ax = \text{proj}_{R(A)(b)}$
  - and $b - Ax \in R(A)^{\perp} = N(A^T)$
  - as well as $A^{T(b - Ax)} = 0$
  - finally $A^T Ax = A^T b$, so either there is a solution or not

if $A^T A$ mxm and let $A^T A \neq 0$ (Why ? exercise)

- Note: cond$(A^T A) =$ rank$(A) \Rightarrow$ could potentially give errors.
- *Theroem* (QR equations)

  The least squares approximation is the solution of this system $R_1 x = Q_1^T b$, furthermore the residual, $\|Ax - b\| = \left\|Q_2^T b\right\|$. We get an upper triangular matrix so it's very easy to solve, and the residual is just measuring the error of our QR equation.

  Proof we know from the projection theorem that $Ax = \text{proj}_{R(A)}(b)$. We also know that $Q_1^T \cdot Q_1 = I$

$$
\begin{aligned}
Ax = \mathrm{proj}_{R(A)}(b) &= Q_1 \cdot Q_1^T b \\
Q_1 \cdot R_1 x &= Q_1 \cdot Q_1^T b \\
Q_1^T \cdot Q_1 R_1 x &= Q_1^T \cdot Q_1 \cdot Q_1^T b \\
R_1 x &= Q_1^T b
\end{aligned}
$$

0.2 Fitting models to Data

- Consider $N + 1$ points $(t_0, y_0), ..., (t_N, y_N)$. Choose a linear model $f(t) = c_0 + c_1 t$. Our job is to find c_0 and c such that $f(t)$ best fits the data.
- Our main job is to minimize the error, we get this formula **PASTE HERE**. We choose the difference to be be squared basically cause the answer is easier to find.