EE445-S22 2022.04.27

Topic: Midterm Review

Lecturer: L.J. Ratliff

# 1 Module 1& 2: Review

**Module 2** :

- Least Squares Approximate Solutions: $\hat{x} = (A^\top A)^{-1}A^\top b$
- Data fitting via feature maps:
  - univariate and multivariate
  - examples of feature maps: piecewise linear, polynomial etc
  - (cross) validation and generalization
- Least squares classification

# 2 Problems

**Problem 1.** (Least Squares.) Consider a least squares problem $Ax = b$ where $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$ with $m > n$. Why is $\hat{x} = (A^\top A)^{-1}A^\top b$ the minimizer of $\|Ax - b\|^2$?
**Solution.** We saw in class the calculus approach to this. As an alternative we can complete the suqare by writing $x = \hat{x} + u$ and simplify:

$$
\begin{aligned}
(Ax-b)^\top(Ax-b) &= (A\hat{x} + Au - b)^\top(A\hat{x} + Au - b) \\
&= (Au)^\top(Au) + (A\hat{x} - b)^\top(A\hat{x} - b) + 2(Au)^\top(A\hat{x} - b) \\
&= \|Au\|^2 + \|A\hat{x} - b\|^2 + 2u^\top \underbrace{(A^\top A\hat{x} - A^\top b)}_{=0 \text{ normal eqns}} \\
&= \|Au\|^2 + \|A\hat{x} - b\|^2 \geq \|A\hat{x} - b\|^2
\end{aligned}
$$

This is clearly minimized by $u = 0$. This is just the Pythagorean theorem, since the residual $r = A\hat{x} - b$ is orthogonal to the space spanned by the columns of $A$—i.e., $0 = A^\top A^\top Ax - A^\top b$.

**Problem 2.** (Least Squares Properties.) Suppose $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. Answer True or False and give a justification:

a. The general least-squares problem is to find a vector, $x$, that makes $Ax$ as close as possible to $b$.
**Solution. True**. By definition, the least squares problem is to find $\hat{x}$ such that $\|A\hat{x} - b\| \leq \|Ax - b\|$ for all $x \in \mathbb{R}^n$. Put in words, this translates to finding a vector $x$ that makes $Ax$ (which is in the col($A$)) as close to $b$ as possible

b. If $b$ is in the column space of $A$, then every solution of $Ax = b$ is a least squares solution.
**Solution. True**. If $b$ is in the column space of $A$ then there is at least one exact solution to the system $Ax = b$. In this case every solution to the system minimizes $\|Ax - b\|$ (makes it zero) and therefore is a least squares solution

c. A least-squares solution of $Ax = b$ is a vector $\hat{x}$ that satisfies $A\hat{x} = \hat{b}$ where $\hat{b}$ is the orthogonal projection of $b$ on to $\text{col}(A) = \text{range}(A)$.
**Solution. True**.
The vector in the column space of $A$ that is closest to $b$ is the orthogonal projection of $b$ in the column space. Then by equivalence to part a. the statement is true.

d. If the columns of $A$ are linearly independent, then the equation $Ax = b$ has exactly one least-squares solution.

**Solution. True.**

By the previous part, the solution to the least squares problem will be the solution of $Ax = \hat{b}$ where $\hat{b}$ is the projection of $b$ in the column space. Then since the columns of $A$ are linearly independent the solution of that system will exist (since the right hand side is in the column space by construction) and it will be unique. Thus the least squares solution will be unique, whenever the columns of $A$ are independent.