Lecture 17, Nov 17, 2021

General Least Squares Continued

- Define a vector space that consists of all the linear combinations of the columns of matrix A
- In a least squares problem $\vec{b}$ does not lie in this space
- We want to find $\vec{x}$ such that $\vec{e} = A\vec{x} - \vec{b}$ is minimized
- The smallest such $\vec{e}$ would be orthogonal to the column space of A since shortest distance is the orthogonal distance
    - This means that $\vec{e} \perp \vec{a}_i$ where $\vec{a}_i$ are the vectors in the columns of the matrix
- Since multiplying a matrix by a vector computes the dot product between the rows of the matrix and the vector, we can compute the dot product between the columns of the matrix and the vector by transposing the matrix
- Therefore $A^T \vec{e} = \vec{0}$ since orthogonality means dot product is 0
- $\quad A^T \vec{e} = \vec{0}$

$$\implies A^T(A\vec{x} - b) = \vec{0}$$

$$\implies A^T A\vec{x} - A^T b = \vec{0}$$

$$\implies A^T A\vec{x} = A^T b$$

- This is called the *normal* system of equations
- Define $A^* = A^T A$ and $\vec{b}^* = A^T \vec{b}$ then $A^* \vec{x} = \vec{b}^*$
- This system does have a solution since $A^T A$ is invertible (but only if A has linearly independent columns)
    * Proof: $A^T A\vec{x} = \vec{0} \implies A\vec{x} \in N(A^T)$, but $A\vec{x} \in C(A)$ and $A\vec{x} \in N(A^T) \implies A\vec{x} = \vec{0}$ since $C(A) \perp N(A^T)$; therefore $\vec{x} \in N(A)$, and if A has linearly independent columns the only such $\vec{x} = \vec{0} \implies A\vec{x} = \vec{0} \implies N(A^T A) = \{\vec{0}\} \implies A^T A$ is invertible
        - Note $N(A) \perp C(A^T)$ because $A\vec{x} = \vec{0}$ requires that the dot product of $\vec{x}$ with every row of A is 0, and thus also the dot product of $\vec{x}$ with all the linear combinations of the rows of A is 0 and so the two subspaces are orthogonal
        - So if we transpose the two we get $N(A^T) \perp C(A)$
    - A^* is a square $n \times n$ matrix where n is the number of unknowns
    - Solving directly for the least squares solution gets us $\vec{x} = (A^T A)^{-1} A^T \vec{b}$