Geometry behind Linear Least Squares

- Let the columns of $A \in R^{m \times n}$ be denoted as $A = [A_1, \ldots, A_n]$ where each $A_i \in R^m$.
  - ◇ The product Ax can be written as
  $$Ax = \sum_{i=1}^{n} x_i A_i,$$
  i.e., Ax is a linear combination of columns of A and hence is an element in the range space of A.
  - ◇ Solving the equation $Ax = b$ is equivalent to finding an appropriate combination of columns of A that makes up the vector b.
  - ◇ A necessary condition for $Ax = b$ to have a solution is that $b \in R(A)$.
- When $b \notin R(A)$, the best we can hope for is to find a combination so that the residual $b - Ax$ is minimized.
  - ◇ In the sense of least squares, the residual $b - Ax$ must be perpendicular to $R(A)$. This geometric setting has rich applications.
- Fredholm Alternative Theorem: Let $A \in R^{m \times n}$ denote a linear transformation from R^n to R^m. Then $R^m = R(A) \oplus N(A^T)$, i.e., for every $z \in R^m$ there exist a unique $y \in R(A)$, and unique $w \in N(A^T)$ such that $z = y + w$.
  - ◇ It suffices to prove $R(A)^{\perp} = N(A^T)$.
  - ◇ $w \in N(A^T) \Leftrightarrow A^T w = 0 \Leftrightarrow x^T A^T w = 0$ for all $x \in R^n \Leftrightarrow w \perp Ax$ for all $x \in R^n \Leftrightarrow w \perp R(A)$.
- Solve the linear least squares problem via the normal equation:
  $$A^T(Ax) = A^T b. \tag{0.1}$$
  - ◇ The normal equation follows from the facts that $b - Ax \perp R(A)$ and hence $A^T(b - Ax) = 0$.