2017-11-28

- For any α, projections of b onto αa are identical.
- Given m linear equations of one variable $a_i x = b_i$, for $i = 1, \dots, m$. Let $\epsilon^2 = \sum(a_i\bar{x} - b_i)^2$.
  - If there exists a solution x s.t. $(a_1, \dots, a_m)x = (b_1, \dots, b_m)$, then $\epsilon = 0$.
  - Otherwise, there is an approximation solution $\bar{x}$ s.t. ϵ^2 is minimized.
  - Let $\frac{de^2}{d\bar{x}} = \sum 2a_i(a_i\bar{x} - b_i) = 0$. Then, $\bar{x} = \frac{a^\top b}{a^\top a}$.
  - $\bar{x}$ is the coefficient s.t. $\bar{x}a$ is the projection of b onto a.
- Given m linear equations of n variables $A_{m\times n}x = b_{m\times 1}$, where $m > n$. Let $\epsilon^2 = \sum(a_i\bar{x} - b_i)^2$, where a_i's are the rows of A for $i = 1, \dots, m$.
  - The error vector $\epsilon = A\bar{x} - b$ is perpendicular to every column of A, i.e., $A^\top(A\bar{x} - b) = 0$. Then, $A^\top Ax = A^\top b$.
  - The sum of square error (SSE) $\epsilon^2 = \|A\bar{x} - b\|^2 = (A\bar{x} - b)^\top(A\bar{x} - b)$. Let $\frac{de^2}{d\bar{x}} = 2A^\top Ax - 2A^\top b = 0$. Then, $A^\top Ax = A^\top b$.
- The **least square solution** to an inconsistent system $Ax = b$ of m equations in n unknowns satisfies $A^\top Ax = A^\top b$, which is referred to as the **normal equations**.
- The properties of $A^\top A$:
  - Every entry of $A^\top A$ is the inner product of the i-th column and j-th column of A.
  - Symmetric. $(A^\top A)^\top = A^\top A$.
  - $A^\top A$ has the same nullspace as A.
    - If $Ax = 0$, then $A^\top Ax = 0$. Hence, $N(A) \subseteq N(A^\top A)$.
    - If $A^\top Ax = 0$, then $x^\top A^\top Ax = (Ax)^\top(Ax) = \|Ax\|^2 = 0$ iff $Ax = 0$. Hence, $N(A^\top A) \subseteq N(A)$.
  - Positive semidefinite.
- *Lemma*: If $A_{m\times n}$ has independent columns, then $A^\top A$ is nonsingular.
  - Rank$(A) = n$ and $N(A) = \{0\}$. Hence, $N(A^\top A) = \{0\}$.
- The least square solution to the inconsistent system $Ax = b$ is the solution of $A^\top Ax = A^\top b$.
  - If the columns of A are linearly independent, then $A^\top A$ is invertible. Hence, $x = (A^\top A)^{-1}A^\top b$.
  - Otherwise, $A^\top A$ is singular and $A^\top Ax = A^\top b$ has infinitely many solutions.
- The normal equation $A^\top Ax = A^\top b$ is always consistent.
- Let A be an $m \times n$ matrix over $\mathbb{R}$. Let $b \notin C(A)$. The closest point to b in $C(A)$ is $p = A(A^\top A)^{-1}A^\top b$. Let $P = A(A^\top A)^{-1}A^\top$.
  - P is the projection matrix that projects any vectors onto $C(A)$.
  - The column space of P is identical to the column space of A, i.e., $C(P) = C(A)$.
- An **orthogonal matrix** Q is a square matrix satisfying $Q^\top Q = I$, i.e., the columns of Q are orthonormal and $Q^{-1} = Q^\top$.
- Examples of orthogonal matrix: rotation matrix, permutation matrix.
- *Proposition*: Q preserves (1) length. $\forall x$, $\|Qx\| = \|x\|$. (2) inner product. $\forall x, y$, $\langle Qx, Qy\rangle = \langle x, y\rangle$. (3) angle. $\forall x, y$, $\angle(x, y) = \angle(Qx, Qy)$.