14 Orthogonal Projection

Orthogonal Projection For $\boldsymbol{v} \in \mathcal{V}$, let $\boldsymbol{v} = \boldsymbol{m} + \boldsymbol{n}$, where $\boldsymbol{m} \in \mathcal{M}$ and $\boldsymbol{n} \in \mathcal{M}^{\perp}$.

- $\boldsymbol{m}$ is called the *orthogonal projection* of $\boldsymbol{v}$ onto $\mathcal{M}$.
- The projector $P_{\mathcal{M}}$ onto $\mathcal{M}$ along $M^{\perp}$ is called the *orthogonal projector* onto $\mathcal{M}$.
- $P_{\mathcal{M}}$ is the unique linear operator such that $P_{\mathcal{M}}\boldsymbol{v} = \boldsymbol{m}$.

Constructing Orthogonal Projectors Let $\mathcal{M}$ be an r-dimensional subspace of $\mathbb{R}^n$, and let the columns of $M_{n\times r}$ and $N_{n\times n-r}$ be bases for $\mathcal{M}$ and $M^{\perp}$, respectively. The orthogonal projectors onto $\mathcal{M}$ and $\mathcal{M}^{\perp}$ are

- $P_{\mathcal{M}} = M(M^{\top}M)^{-1}M^{\top}$ and $P_{\mathcal{M}^{\perp}} = N(N^{\top}N)^{-1}N^{\top}$.

If $\mathcal{M}$ and $\mathcal{N}$ contain orthonormal bases for $\mathcal{M}$ and $\mathcal{M}^{\perp}$, then

- $P_{\mathcal{M}} = MM^{\top}$ and $P_{\mathcal{M}^{\perp}} = NN^{\perp}$.
- $P_{\mathcal{M}} = U\begin{pmatrix} I_{r\times r} & \mathbf{0} \\ \mathbf{0} & \mathbf{0}\end{pmatrix}U^{\top}$, where $U = (M|N)$.
- $P_{\mathcal{M}^{\perp}} = I - P_{\mathcal{M}}$ in all cases.

Matrix 2-norm[12]

Orthogonal Projectors Suppose that $P \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ is a projector - i.e., $P^2 = P$. The following statements are equivalent to saying that P is an *orthogonal* projector.

- $\mathrm{im}(P)\perp\ker(P)$.
- $P^{\top} = P$ ((i.e., orthogonal projector $\Leftrightarrow P^2 = P = P^{\top}$).
- $\|P\|_2 = 1$ for the matrix 2-norm.

1. Let $\boldsymbol{u} \in \mathbb{R}^n$, $\boldsymbol{u} \neq \mathbf{0}$ and consider the line $\mathcal{L} = \mathrm{span}\{\boldsymbol{u}\}$. Construct the orthogonal projector onto $\mathcal{L}$, and then determine the orthogonal projection of a vector $\boldsymbol{x} \in \mathbb{R}^n$ onto $\mathcal{L}$.

2. For $A \in \mathrm{Mat}_{m\times n}$ such that $\mathrm{rank}(A) = r$, describe the orthogonal projectors onto each of the four fundamental subspaces of A.

Closest Point Theorem Let $\mathcal{M}$ be a subspace of an inner-product space $\mathcal{V}$, and let $\boldsymbol{b}$ be a vector in $\mathcal{V}$. The unique vector in $\mathcal{M}$ that is closest to $\boldsymbol{b}$ is $\boldsymbol{p} = P_{\mathcal{M}}\boldsymbol{b}$, the orthogonal projection of $\boldsymbol{b}$ onto $\mathcal{M}$. In other words,

$$\min_{\boldsymbol{m}\in\mathcal{M}} \|\boldsymbol{b} - \boldsymbol{m}\|_2 = \|\boldsymbol{b} - P_{\mathcal{M}}\boldsymbol{b}\|_2 = \mathrm{dist}(\boldsymbol{b}, \mathcal{M}).$$

This is called the *orthogonal distance* between $\boldsymbol{b}$ and $\mathcal{M}$.

3. Find the orthogonal projection of $\boldsymbol{b}$ onto $\mathcal{M} = \mathrm{span}\{u\}$, and then determine the orthogonal projection of $\boldsymbol{b}$ onto $M^{\perp}$, where $\boldsymbol{b} = (4, 8)^{\top}$ and $\boldsymbol{u} = (3, 1)^{\top}$.

4. Let $A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 4 & 1 \\ 1 & 2 & 0 \end{pmatrix}$ and $\boldsymbol{b} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. (a) Compute the orthogonal projectors onto each of the four fundamental subspaces associated with A. (b) Find the point in $\ker(A)^{\perp}$ that is closest to $\boldsymbol{b}$.

5. For an orthogonal projector P, prove that $\|Px\|_2 = \|x\|_2$ if and only if $x \in \mathrm{im}(P)$.

6. Explain why $A^{\top}P_{\mathrm{im}(A)} = A^{\top}$ for all $A \in \mathrm{Mat}_{m\times n}$.

7. Explain why $P_{\mathcal{M}} = \sum_{i=1}^{r} \boldsymbol{u}_i\boldsymbol{u}_i^{\top}$ whenever $\mathcal{B} = \{\boldsymbol{u}_1, \boldsymbol{u}_2, ..., \boldsymbol{u}_r\}$ is an orthonormal basis for $\mathcal{M} \subseteq \mathbb{R}^n$.

8. Explain how to use orthogonal reduction techniques to compute the orthogonal projectors onto each of the four fundamental subspaces of a matrix $A \in \mathrm{Mat}_{m\times n}$.

9. Describe all 2×2 projectors in $\mathrm{Mat}_{2\times 2}(\mathbb{R})$.

10. The line $\mathcal{L}$ in $\mathbb{R}^n$ passing through two distinct points $\boldsymbol{u}$ and $\boldsymbol{v}$ is $\mathcal{L} = \boldsymbol{u} + \mathrm{span}\{\boldsymbol{u} - \boldsymbol{v}\}$. If $\boldsymbol{u} \neq \mathbf{0}$ and $\boldsymbol{v} = \alpha\boldsymbol{u}$, then $\mathcal{L}$ is a line not passing through the origin - i.e., $\mathcal{L}$ is not a subspace. Sketch a picture in $\mathbb{R}^2$ or $\mathbb{R}^3$ to visualize this, and then explain how to project a vector $\boldsymbol{b}$ orthogonally onto $\mathcal{L}$.

[12] The matrix norm induced by the euclidean vector norm is

$$\|A\|_2 = \max_{\|x\|_2=1} \|A\boldsymbol{x}\|_2 = \sqrt{\lambda_{max}},$$

where λ_{max} is the largest number λ such that $A^*A - \lambda I$ is singular. In case when A is nonsingular,

$$\|A^{-1}\|_2 = \frac{1}{\min\{\|A\boldsymbol{x}\|_2 \,:\, \|x\|_2 = 1\}} = \frac{1}{\sqrt{\lambda_{min}}},$$

where λ_{min} is the smallest number λ such that $A^*A - \lambda I$ is singular. If you are already familiar with eigenvalues, these say that λ_{max} and λ_{min} are the largest and smallest eigenvalues of A^*A.

Classical Least Squares[13]

Least Squares Solutions Each of the following four statements is equivalent to saying that $\widehat{x}$ is a least squares solution for a possibly inconsistent linear system $A\boldsymbol{x} = \boldsymbol{b}$.

- $\|A\widehat{\boldsymbol{x}} - \boldsymbol{b}\| = \min_{\boldsymbol{x}\in\mathbb{R}^n} \|A\boldsymbol{x} - \boldsymbol{b}\|_2$
- $A\widehat{\boldsymbol{x}} = P_{\mathrm{im}(A)}\boldsymbol{b}$.
- $A^\top A\widehat{\boldsymbol{x}} = A^\top \boldsymbol{b}$
  ($A^* A\widehat{x} = A^*\boldsymbol{b}$ where $A \in \mathrm{Mat}_{m\times n}(\mathbb{C})$).
- $\widehat{x} \in A^\star \boldsymbol{b} + \ker(A)$
  ($A^\star\boldsymbol{b}$ is the minimal 2-norm LSS).

Caution! These are valuable theoretical characterizations, but none is recommended for floating-point computation. Directly solving second of third from above or explicitly computing $A^\star$ can be inefficient and numerically unstable.

11. Let $\mathcal{P}_2$ denote a vector space of all polynomial of degree ≤ 2, $\mathcal{P}_2 = \{ax^2 + bx + c \,|\, a, b, c \in \mathbb{R}\}$.

(a) Check is it $\langle p, q\rangle = p(1)q(1) + 2p(0)q(0) + p(-1)q(-1)$ inner product for $\mathcal{P}_2$.

(b) For subspace $\mathcal{L} \subseteq \mathcal{P}_2$ generated by $p_1(x) = 1$ and $p_2(x) = x$ find an orthogonal complement.

(c) Find the orthogonal projection of $p(x) = -2x^2 + x + 2$ on $\mathcal{L}$.

12. In inner product space $\mathbb{R}^4$, with standard inner product, let $\mathcal{M}$ denote subspace spanned by vectors $(2, 1, 0, 0)^\top$ and $(1, 1, 1, 1)^\top$. Find a basis for orthogonal complement of $\mathcal{M}$ and find the orthogonal projection of $\boldsymbol{a} = (3, -4, 5, -5)^\top$ onto $\mathcal{M}$.

13. Let $\mathcal{M} = \mathrm{span}\{\boldsymbol{a}, \boldsymbol{b}\}$ denote subspace of inner product space $\mathbb{R}^n$ (with standard inner product) spaned by vectors $\boldsymbol{a} = (0, 1, 2, ..., n-1)^\top$ and $\boldsymbol{b} = (1, 1, 1, ..., 1)^\top$. Find orthogonal complement $\mathcal{M}^\perp$ and find the orthogonal projection of $\boldsymbol{z}$ onto $\mathcal{M}$ where

$$\boldsymbol{z} = \left(\frac{1}{2}n(3-n), \frac{1}{2}n(n-1), 0, 0, ..., 0\right)^\top \in \mathbb{R}^n.$$

14. Find the orthogonal projection of $\boldsymbol{x} = (-12, -13, 5, 2)^\top$ onto $\mathcal{M}$ if we have that

$$\mathcal{M} = \mathrm{span}\left\{\begin{pmatrix}1\\-2\\2\\-3\end{pmatrix}, \begin{pmatrix}2\\-3\\2\\4\end{pmatrix}\right\} \subseteq \mathbb{R}^4$$

(with respect to standard inner product).

15. In inner product space $\mathcal{P}_3 = \{at^3 + bt^2 + ct + d \,|\, a, b, c, d \in \mathbb{R}\}$ of all polynomials of degree ≤ 3 with an inner product

$$\langle p, q\rangle = \int_{-1}^{1} p(t)q(t)\,\mathrm{d}t$$

let $\mathcal{M} = \mathrm{span}\{t, 1+t\}$ be a given subspace. Find the orthogonal projection of $r(t) = -5t^3 - 12t^2 + 6t + 6$ onto $\mathcal{M}$.

16. Space $\mathcal{L}$ is defined as set of solutions for the following system

$$\begin{aligned} 2x_1 + x_2 + x_3 + 3x_4 &= 0\\ 3x_1 + 2x_2 + 2x_3 + x_4 &= 0\\ x_1 + 2x_2 + 2x_3 - 9x_4 &= 0\ . \end{aligned}$$

Find the orthogonal projection of $x = (7, -4, -1, 2)^\top$ onto $\mathcal{L}$ in $\mathbb{R}^4$.

[13]The following problem arises in almost all areas where mathematics is applied. At discrete points t_i (often points in time), observations b_i of some phenomenon are made, and the results are recorded as a set of ordered pairs

$$\mathcal{D} = \{(t_1, b_1), (t_2, b_2), ..., (t_m, b_m)\}.$$

On the basis of these observations, the problem is to make estimations or predictions at points (times) $\widehat{t}$ that are between or beyond the observation points t_i. A standard approach is to find the equation of a curve $y = f(t)$ that closely fits the points in $\mathcal{D}$ so that the phenomenon can be estimated at any nonobservation point $\widehat{t}$ with the value $\widehat{y} = f(\widehat{t})$.

General Least Squares Problem is the following. For $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ and $b \in \mathbb{R}^m$, let $\varepsilon = \varepsilon(x) = A\boldsymbol{x} - \boldsymbol{b}$. The general least squares problem is to find a vector $\boldsymbol{x}$ that minimizes the quantity

$$\sum_{i=1}^{m} \varepsilon_i^2 = \varepsilon^\top\varepsilon = (A\boldsymbol{x} - \boldsymbol{b})^\top(A\boldsymbol{x} - \boldsymbol{b}).$$

Any vector that provides a minimum value for this expression is called a **least squares solution**.

- The set of all least squares solutions is precisely the set of solutions to the system of normal equations $A^\top A\boldsymbol{x} = A^\top\boldsymbol{b}$.
- There is a unique least squares solution if and only if $\mathrm{rank}(A) = n$, in which case it is given by $\boldsymbol{x} = (A^\top A)^{-1}A^\top\boldsymbol{b}$.
- If $A\boldsymbol{x} = \boldsymbol{b}$ is consistent, then the solution set for $A\boldsymbol{x} = \boldsymbol{b}$ is the same as the set of least squares solutions.