MODULE 1.3
Orthogonal Projections in R^m

by
S.Lakshmivarahan
School of Computer Science
University of Oklahoma
Norman, OK-73019, USA
varahan@ou.edu

- Let $x, y \in R^m$. The inner product is defined as

$$< x, y >= x^T y = \sum_{i=1}^{n} x_i y_i = y^T x =< y, x > \tag{1}$$

- Norm of a vector x is given by

$$||x|| =< x, x >^{\frac{1}{2}} = (\sum_{i=1}^{m} x_i^2)^{\frac{1}{2}} \tag{2}$$

- Cauchy- Schwartz inequality: From

$$< x, y >= ||x|| \, ||y|| \, cos(\theta) \tag{3}$$

it follows that

$$|cos\theta| = \frac{| < x, y > |}{||x|| \, ||y||} \leq 1 \tag{4}$$

Projection of y along x - a geometric view

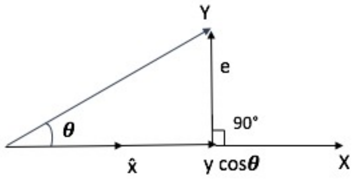
- Let $\hat{x} = \frac{x}{||x||}$ be the unit vector,

$$||\hat{x}|| = 1 \tag{5}$$

- $$< y, \hat{x} >= y^T \hat{x} = ||y|| cos(\theta) \tag{6}$$

  which is the component of y in the direction $\hat{x}$.

- The vector $(ycos\theta)\hat{x}$ is called the projection of y along $\hat{x}$



Orthogonality of this projection

- Let $e = y - (ycos\theta)\hat{x}$ be the error in the projection
- Then,

$$< e, \hat{x} >= e^T\hat{x} = y^T\hat{x} - (ycos\theta)\hat{x}^T\hat{x} = 0 \tag{7}$$

- Hence the name orthogonal projection

Analytical expression for orthogonal projection

- Let $h \in R^m$ and $x \in R^m$ be any other vector
- Any vector along h can be expressed as a multiple $h\alpha$ for some real α
- Problem: Given x and h, find $\alpha \in R$ that minimizes the distance between x and $h\alpha$
- That is, find α that minimizes

$$\begin{aligned} Q(\alpha) = \|x - h\alpha\|^2 &= (x - h\alpha)^T(x - h\alpha) \\ &= x^T x - 2x^T h\alpha + \alpha^2 h^T h \end{aligned} \tag{8}$$

Optimal α

- Minimizer α^* is obtained by solving

$$0 = \frac{dQ}{d\alpha} = -2h^T x + 2\alpha h^T h \tag{9}$$

- That is,

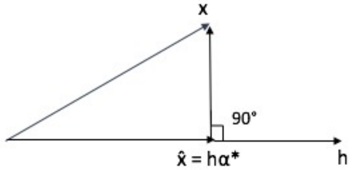
$$\alpha^* = (h^T h)^{-1} h^T x = h^+ x \tag{10}$$

where

$$h^+ = (h^T h)^{-1} h^T \tag{11}$$

is called the generalized inverse of h

- The orthogonal projection of x along h is given by



$$\hat{x} = h\alpha^* = hh^+x = h(h^Th)^{-1}h^Tx = P_hx \tag{12}$$

where

$$P_h = hh^+ = h(h^Th)^{-1}h^T \tag{13}$$

is called the orthogonal projection matrix

- Let
$$e = x - \hat{x} = (I - P_h)x \qquad (14)$$
be the error in the projection
- Clearly:
$$h^T e = (h^T - h^T P_h)x = 0 \qquad (15)$$
since $h^T P_h = (h^T h)(h^T h)^{-1} h^T = h^T$
- Hence, P_h is called the orthogonal projection operator

- Symmetry: $P_h^T = P_h$
- Idempotent: $P_h^2 = P_h$
- P_h is a rank one matrix
- $det(P_h) = 0$, that is, P_h is singular
- 1 is the only non-zero eigenvalue of P_h
- P_h is not an orthogonal matrix: $P_h^T \neq P_h^{-1}$ since P_h^{-1} is not defined

A Generalization

- Let $H \in R^{m\times n}$ with m>n and Rank(H) = n
- Then, $(H^T H) \in R^{n\times n}$ is SPD
- Let $x \in R^m$
- Problem: Find an $\alpha \in R^n$ such that $\hat{x} = H\alpha \in R$ and

$$\begin{aligned} Q(\alpha) &= (x - H\alpha)^T(x - H\alpha) = ||x - H\alpha||^2 \\ &= x^T x - 2x^T H\alpha + \alpha^T(H^T H)\alpha \end{aligned} \tag{16}$$

is a minimum

- From
$$\nabla_\alpha Q(\alpha) = -2H^T x + 2(H^T H)\alpha = 0 \tag{17}$$
it follows that
$$\alpha^* = (H^T H)^{-1} H^T x = H^+ x \tag{18}$$
minimizes $Q(\alpha)$ since
$$\nabla_\alpha^2 Q(\alpha) = (H^T H) \quad is \quad SPD \tag{19}$$
- $H^+ = (H^T H)^{-1} H^T \in R^{n \times m}$ is called the generalized inverse of H

Optimal projection

- Then

$$\hat{x} = H\alpha^* = H(H^T H)^{-1} H^T x = HH^+ x = P_H x \tag{20}$$

  where

$$P_H = H(H^T H)^{-1} H^T \in R^{m\times m} \tag{21}$$

  is the projection operator in R^m onto the n-dimensional subspace spanned by the columns of H

- $$e = x - \hat{x} = (I - P_H)x \tag{22}$$

  is the error in this projection

- Verify that

$$e^T H = 0 \tag{23}$$

  and hence the name orthogonal projection

Properties of P_H

- Symmetry: $P_H^T = P_H$
- Idempotent: $P_H^2 = P_H$
- $RANK(P_H) = n$ since that of H is n
- P_H is singular
- There are exactly n non-zero eigenvalues of P_H
- P_H is not an orthogonal matrix

Reference

- Chapters 5 and 6 in J. Lewis, S. Lakshmivarahan and S.K. Dhall (2006) Dynamic Data Assimilation, Cambridge University Press.

Exercises

1) Let $x \in R^m$ and $h = (1, 1, \ldots, 1)^T \in R^m$ and $h = (1, 1, 1, \ldots, 1)^T \in R^m$ be a vector all of whose components are 1. Compute an expression for α that minimizes the distance between x and $h\alpha$.

2) Let H = $\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}$

- Compute $H^T H$, HH^T, H^+, P_H, HH^+, $H^+ H$, $HH^+ H$, $H^+ HH^+$
- Compute the eigenvalues of P_H

3) Verify the following:
a) $HH^+H = H$
b) $H^+HH^+ = H^+$
c) $(H^+H)^T = H^+H$
d) $(HH^+)^T = HH^+$
Note: Any H^+ satisfying the properties (a)-(d) is called the Moore-Penrose inverse.

4) Given P_H, define $P_H^\perp = I - P_H$
- Verify that $P_H^\perp$ is symmetric and idempotent
- For the H in problem 2, Compute $P_H^\perp$ and its rank