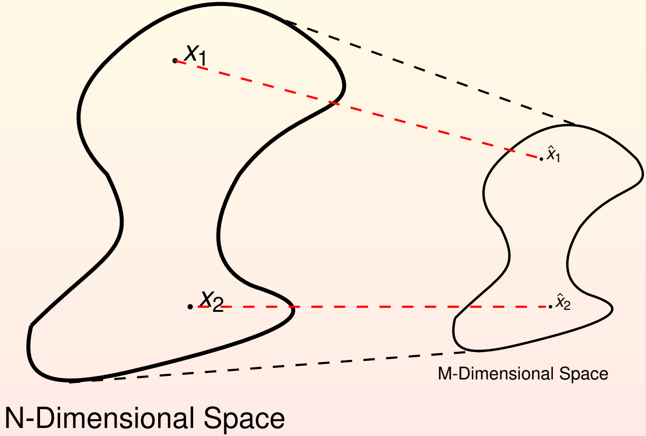
Principal Components

Robert M. Haralick

Computer Science, Graduate Center
City University of New York

Space Squeezing: Dimensionality Reduction



Introduction

- Assume you know linear algebra
- Present concepts in a slightly different way
- Not a formal presentation of linear algebra
- Will introduce projection operators in a formal way

Spaces

- Spaces have points called vectors
- Spaces have sets of points
- Some sets are called subspaces
- Spaces have directions
- Spaces have sets of directions
- Spaces have a language of representing points in terms of traveling different lengths in different directions

Language of Spaces

- The words of the language are the basis elements
  - $b_1, b_2, \ldots, b_N, ||b_n|| = 1, n = 1, \ldots, N$
- The basis elements specify independent directions
- The sentence takes the form $\sum_{n=1}^{N} \alpha_n b_n$
- The meaning of the sentence is
  - Begin at the origin
  - Go α_1 in direction b_1
  - Go α_2 in direction b_2
  - $\ldots$
  - Go α_N in direction b_N
  - And you arrive at the point represented by $(\alpha_1, \ldots, \alpha_N)$
- The set of all places that can be reached by such a sentence is called the space spanned by the directions $b_1, \ldots, b_N$
- The interesting sentences are the minimal ones

Linear Independence

Minimal sentence means using independent directions.

Definition

$b_1, \ldots, b_N$ are independent directions (linearly independent) when

$$\sum_{n=1}^{N} \alpha_n b_n = 0 \text{ if and only if } \alpha_n = 0, n = 1, \ldots, N$$

If you travel α_1 in direction b_1, then travel α_2 in direction b_2, ..., then travel α_N in direction b_N and you return to the origin, then the directions are dependent.

Linear Dependence

Definition

$b_1, \ldots, b_N$ are linearly dependent if and only if for some $\alpha_1, \ldots, \alpha_N$, not all 0

$$\sum_{n=1}^{N} \alpha_n b_n = 0$$

If you travel α_1 in direction b_1, then travel α_2 in direction b_2, ..., then travel α_N in direction b_N and you return to the origin, then the directions are dependent.

Angles and Inner Products

- $x'y$ is the inner product of x and y
- $b_1, \ldots, b_N$ set of norm 1 basis vectors
- $b_i' b_i = 1$, norm 1
- $b_i' b_j$ cosine of the angle between directions b_i and b_j
- $b_i' b_j = b_j' b_i$
- $b_i' b_i = ||b_i||^2$
- $b'(c + d) = b'c + b'd$
- $(\alpha b)'c = \alpha(b'c)$
- $b_i \perp b_j$ geometrically orthogonal
- $b_i \perp b_j$ if and only if $b_i' b_j = 0$
- $b_1, \ldots, b_N$ is orthonormal if and only if
  - $b_i' b_j = 0$ when $i \neq j$
  - $||b_i|| = 1$

Lengths

Definition

The length of a vector x is its distance from the origin.

$$||x|| = \sqrt{x'x}$$

Let $x = \sum_{n=1}^{N} \alpha_n b_n$ and $b_1, \ldots, b_N$ be orthonormal

$$\begin{aligned}
||x||^2 &= ||\sum_{i=1}^{N} \alpha_i b_i||^2 = (\sum_{i=1}^{N} \alpha_i b_i)' \sum_{j=1}^{N} \alpha_j b_j \\
&= \sum_{i=1}^{N} \alpha_i b_i' (\sum_{j=1}^{N} \alpha_j b_j) = \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j b_i' b_j \\
&= \sum_{i=1}^{N} \alpha_i^2 b_i' b_i + \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} \alpha_i \alpha_j b_i' b_j = \sum_{i=1}^{N} \alpha_i^2
\end{aligned}$$

Coordinate Representation

- Let $b_1, \ldots, b_N$ be an orthonormal basis
- Let x be a vector
- Find $\alpha_1, \ldots, \alpha_N$ such that $x = \sum_{n=1}^{N} \alpha_n b_n$

Suppose $x = \sum_{n=1}^{N} \alpha_n b_n$.

$$
\begin{aligned}
b_i' x &= b_i' \left(\sum_{n=1}^{N} \alpha_n b_n\right) \\
&= \sum_{n=1}^{N} \alpha_n b_i' b_n \\
&= \alpha_i b_i' b_i = \alpha_i
\end{aligned}
$$

$x = (\alpha_1, \ldots, \alpha_N)$ with respect to basis $b_1, \ldots, b_n$
Change the basis and you change the coordinate representation.

Inner Product

Let $x = (\alpha_1, \ldots, \alpha_N)$, $y = (\beta_1, \ldots, \beta_N)$, be coordinates with respect to orthonormal basis $b_1, \ldots, b_N$

$$
\begin{aligned}
x'y &= (\sum_{i=1}^{N} \alpha_i b_i)'(\sum_{j=1}^{N} \beta_j b_j) \\
&= \sum_{i=1}^{N} \alpha_i b_i'(\sum_{j=1}^{N} \beta_j b_j) = \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \beta_j b_i' b_j \\
&= \sum_{i=1}^{N} \alpha_i \beta_i b_i' b_i + \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} \alpha_i \beta_j b_i' b_j \\
&= \sum_{i=1}^{N} \alpha_i \beta_i
\end{aligned}
$$

Dimensionality

- N: Dimension of Space
  - Number of directions required in a minimal sentence to specify (reach) any point in the space
  - Number of degrees of freedom needed to represent a point in the space
  - $\{x \mid x = \sum_{n=1}^{N} \alpha_n b_n\}$
- M: Dimension of Subspace
  - Number of directions required in a minimal sentence to specify (reach) any point in the subspace
  - Number of degrees of freedom needed to represent a point in the subspace
  - $\{x \mid x = \sum_{m=1}^{M} \beta_m b_m\}$
  - $\{x \mid x = \sum_{m=1}^{M} \beta_m b_m + \sum_{m=M+1}^{N} 0 b_m\}$
  - M degrees of freedom; $N - M$ degrees of constraint

Constraints

- M: Dimension of Subspace
  - Number of directions required in a minimal sentence to specify (reach) any point in the subspace
  - Number of degrees of freedom needed to represent a point in the subspace
  - $\{x \mid x = \sum_{m=1}^{M} \beta_m b_m\}$
  - $\{x \mid x = \sum_{m=1}^{M} \beta_m b_m + \sum_{m=M+1}^{N} 0 b_m\}$
  - M degrees of freedom; $N - M$ degrees of constraint

Let $i \in \{M+1, \ldots, N\}$ and $b_1, \ldots, b_N$ be orthonormal.
Consider $b_i' x$

$$
\begin{aligned}
b_i' x &= b_i' \sum_{m=1}^{M} \beta_m b_m = \sum_{m=1}^{M} b_i' \beta_m b_m = \sum_{m=1}^{M} \beta_m b_i' b_m \\
&= \sum_{m=1}^{M} \beta_m 0 = 0
\end{aligned}
$$

Co-Dimension

- M: Dimension of Subspace
  - Number of directions required in a minimal sentence to specify (reach) any point in the subspace
  - Number of degrees of freedom needed to represent a point in the subspace
  - $\{x \mid x = \sum_{m=1}^{M} \beta_m b_m\}$
  - $\{x \mid x = \sum_{m=1}^{M} \beta_m b_m + \sum_{m=M+1}^{N} 0 b_m\}$
  - M degrees of freedom
  - $N - M$ degrees of constraint
  - $N - M$ Co-dimension

$N - M$ Constraints

Let $i \in \{M+1, \ldots, N\}$ and $b_1, \ldots, b_N$ be orthonormal.

$$b_i' x = 0, \; i \in \{M+1, \ldots, N\}$$

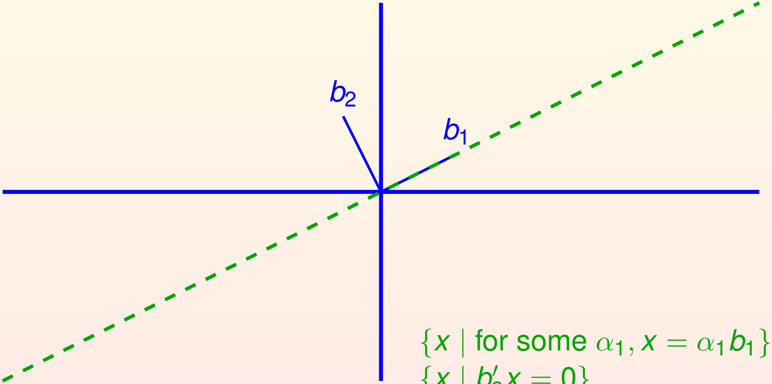
Basis Vectors

Let $b_1, \ldots, b_N$ be an orthonormal basis for a space S.

- Each b_n is a direction
- The length of each b_n is one
- A direction and a length represents a point or vector in the space
- $b_i \perp b_j$, $i \neq j$
- b_n is a point or vector in the space

Representing Subspaces

1-Dimensional Space



$$\{x \mid \text{for some } \alpha_1, x = \alpha_1 b_1\}$$
$$\{x \mid b_2' x = 0\}$$

2-Dimensional Space

Representing Subspaces

- N Dimensional Space S
- M Dimensional Subspace T
- $b_1, \ldots, b_N$ orthonormal basis
- $S = \{x \mid \text{for some } \alpha_1, \ldots, \alpha_N,\ x = \sum_{n=1}^{N} \alpha_n b_n\}$
- M degrees of freedom
  - $T = \{x \mid \text{for some } \alpha_1, \ldots, \alpha_M,\ x = \sum_{m=1}^{M} \alpha_m b_m\}$
- $N - M$ degrees of constraint
  - $T = \{x \mid b_i' x = 0, i \in \{M+1, \ldots, N\}\}$

Orthogonal Subspaces

Definition

A subspace T is orthogonal to a subspace U if and only if $t \in T$ and $u \in U$ implies

$$t'u = 0$$

Definition

Let T be a subspace of S. The orthogonal complement of T, denoted by $T^\perp$, is defined by

$$T^\perp = \{x \in S \mid \text{for every } t \in T,\ x't = 0\}$$

Orthogonal Subspaces

Proposition

Let $b_1, \dots, b_N$ be an orthonormal basis of S. Let V be a subspace of S spanned by $b_1, \dots, b_M$. Then $V^\perp$ is the subspace spanned by $b_{M+1} \dots, b_N$

Proof.

$V^\perp = \{x \in S \mid v \in V \text{ implies } x'v = 0\}$

$$\begin{aligned} x'v &= x' \sum_{m=1}^{M} \alpha_m b_m = (\sum_{n=1}^{N} \beta_n b_n)' \sum_{m=1}^{M} \alpha_m b_m \\ &= \sum_{n=1}^{N} \beta_n \sum_{m=1}^{M} \alpha_m b_n' b_m = \sum_{m=1}^{M} \alpha_m \beta_m \end{aligned}$$

$\sum_{m=1}^{M} \alpha_m \beta_m = 0$ *for all* $\alpha_1, \dots, \alpha_M$ *implies* $\beta_1 = 0, \dots, \beta_M = 0$
Therefore,

$$V^\perp = \{x \mid x = \sum_{i=M+1}^{N} \beta_i b_i\}$$

□

Orthogonal Representations

Proposition

Let V be a subspace of S and let $x \in S$. Then there exists a $v \in V$ and $w \in V^{\perp}$ such that $x = v + w$

Proof.

Let $b_1, \ldots, b_N$ be an orthonormal basis for S such that $b_1, \ldots, b_M$ is an orthonormal basis for V and $b_{M+1}, \ldots b_N$ is an orthonormal basis for $V^{\perp}$ Then for some $\alpha_1, \ldots, \alpha_N$,

$$x = \sum_{n=1}^{N} \alpha_n b_n = \sum_{n=1}^{M} \alpha_n b_n + \sum_{i=M+1}^{N} \alpha_i b_i$$

But $v = \sum_{n=1}^{M} \alpha_n b_n \in V$ and $w = \sum_{i=M+1}^{N} \alpha_i b_i \in V^{\perp}$. Therefore $x = v + w$ for $v \in V$ and $w \in V^{\perp}$. □

Orthogonal Projection

Definition

Let V be a subspace of S. Let $x \in S$ and $x = v + w$ where $v \in V$ and $w \in V^{\perp}$. Then v is called the orthogonal projection of x onto V.

Orthogonal Projections are Unique

Proposition

Let V be a subspace of S. Let $x \in S$ and $x = v_1 + w_1 = v_2 + w_2$ where $v_1, v_2 \in V$ and $w_1, w_2 \in V^\perp$. Then $v_1 = v_2$.

Proof.

Let $b_1, \ldots, b_M$ be an orthonormal basis for V. Then $v_1 = \sum_{m=1}^{M} \alpha_m b_m$ and $v_2 = \sum_{m=1}^{M} \beta_m b_m$.

$$\begin{aligned} b_i' x &= b_i'(v_1 + w_1) = b_i' \sum_{m=1}^{M} \alpha_m b_m = \alpha_i \\ &= b_i'(v_2 + w_2) = b_i' \sum_{m=1}^{M} \beta_m b_m = \beta_i \end{aligned}$$

Therefore, $\alpha_i = \beta_i$, $i = 1, \ldots, M$ □

Orthogonal Projection Operator

Proposition

Let V be an M dimensional subspace of S. Let $x \in S$ and $x = v + w$ where $v \in V$ and $w \in V^{\perp}$. Let $b_1, \ldots, b_N$ be an orthonormal basis of S and $b_1, \ldots, b_M$ be an orthonormal basis of V. Then $v = Px$ where $P = \sum_{m=1}^{M} b_m b_m'$.

Proof.

$x \in S$ implies $x = \sum_{n=1}^{N} \beta_n b_n = \sum_{m=1}^{M} \beta_m b_m + \sum_{n=M+1}^{N} \beta_n b_n$. Then

$$b_m' x = b_m' \sum_{n=1}^{N} \beta_n b_n = \sum_{n=1}^{N} \beta_n b_m' b_n = \beta_m$$

Now,

$$\begin{aligned} v &= \sum_{m=1}^{M} \beta_m b_m = \sum_{m=1}^{M} (b_m' x) b_m = (\sum_{m=1}^{M} b_m b_m') x \\ &= Px \end{aligned}$$

Projection Operators

Definition

P is called a projection operator if and only if $P^2 = P$

$$\begin{pmatrix} .3 & .7 \\ .3 & .7 \end{pmatrix}\begin{pmatrix} .3 & .7 \\ .3 & .7 \end{pmatrix} = \begin{pmatrix} .3 & .7 \\ .3 & .7 \end{pmatrix}$$

$$\begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}$$

$$\begin{pmatrix} .2 & .4 \\ .4 & .8 \end{pmatrix}\begin{pmatrix} .2 & .4 \\ .4 & .8 \end{pmatrix} = \begin{pmatrix} .2 & .4 \\ .4 & .8 \end{pmatrix}$$

Orthogonal Projection Operator

Definition

Let $b_1, \ldots, b_N$ be an orthonormal basis for S and $b_1, \ldots, b_M$ an orthonormal basis for the subspace V of S. Then $P = \sum_{m=1}^{M} b_m b_m'$ is the orthogonal projection operator to V.

Orthogonal Projection Operator

If $b_1, \ldots, b_M$ is an orthonormal basis for a subspace V of S, then the orthogonal projection operator onto V has the following representation.

$$
\begin{aligned}
P &= \sum_{m=1}^{M} b_m b_m' \\
&= \begin{pmatrix} \vdots & \vdots & \ldots & \vdots \\ b_1 & b_2 & \ldots & b_M \\ \vdots & \vdots & \ldots & \vdots \end{pmatrix} \begin{pmatrix} \ldots & b_1' & \ldots \\ \ldots & b_2' & \ldots \\ \vdots & \vdots & \vdots \\ \ldots & b_M' & \ldots \end{pmatrix}
\end{aligned}
$$

Orthogonal Projection Operators

Proposition

If P is an orthogonal projection operator to the subspace V of S, then

- $P^2 = P$
- $P = P'$

Proof.

Let $b_1, \ldots, b_M$ be an orthonormal basis for V. Then,

$$\begin{aligned}
P^2 &= \sum_{m=1}^{M} b_m b_m' \sum_{i=1}^{M} b_i b_i' \\
&= \sum_{m=1}^{M} b_m \sum_{i=1}^{M} (b_m' b_i) b_i' = \sum_{m=1}^{M} b_m b_m' = P \\
P' &= (\sum_{m=1}^{M} b_m b_m')' = \sum_{m=1}^{M} (b_m b_m')' \\
&= \sum_{m=1}^{M} b_m b_m' = P
\end{aligned}$$

Uniqueness

Proposition

Suppose $P = P^2$, $P = P'$, $Q = Q^2$, and $Q = Q'$. If $PQ = Q$ and $QP = P$, then $Q = P$.

Proof.

$$Q = PQ = (PQ)' = Q'P' = QP = P$$

□

Orthogonal Projection Operators are Unique

Proposition

Let V be a M dimensional subspace of S. Let $b_1, \ldots, b_M$ be a basis for V and let $c_1, \ldots, c_M$ be an orthonormal basis for V. Define $P = \sum_{m=1}^{M} b_m b_m'$ and $Q = \sum_{m=1}^{M} c_m c_m'$. Then $Q = P$.

Proof.

By the definition of orthogonal projection operators, both P and Q are orthogonal projection operators onto V. Hence, $P = P^2$ and $P = P'$. Likewise, $Q = Q^2$ and $Q = Q'$. Since the columns of P and Q are in V, $PQ = Q$ and $QP = Q$ By the uniqueness proposition, $Q = P$. □

Orthogonal Projection Operator Characterization Theorem

Theorem

If $P = P^2$ and $P = P'$, then P is the orthogonal projection operator onto $Col(P)$.

Proof.

Let $b_1, \ldots, b_M$ be an orthonormal basis for $Col(P)$. Define $Q = \sum_{m=1}^{M} b_m b_m'$. Then $Q = Q^2$ and $Q = Q'$. Clearly, $Col(Q) = Col(P)$ so that $QP = P$ and $PQ = Q$. By the uniqueness proposition, $P = Q$. And since Q is the orthogonal projection operator onto $Col(P)$, P must also be the orthogonal projection operator onto $Col(P)$. □

Orthogonal Projection Minimizes Error

Theorem

Let V be a subspace of S. Let $f : S \to V$ and $x \in S$.

$$\min_f (x - f(x))'(x - f(x))$$

is achieved when f is the orthogonal projection operator from S to V

Proof.

Let $x \in S$. Then there exists $v \in V$ and $w \in V^{\perp}$ such that $x = v + w$. Consider

$$\begin{aligned}
\epsilon^2 &= (x - f(x))'(x - f(x)) \\
&= x'x - (v + w)'f(x) - f(x)'(v + w) + f(x)'f(x) \\
&= x'x - v'f(x) - f(x)'v - f(x)'f(x) \\
&= (v + w)'(v + w) - v'f(x) - f(x)'v - f(x)'f(x) \\
&= v'v - v'f(x) - f(x)'v + f(x)'f(x) + w'w \\
&= (v - f(x))'(v - f(x)) + w'w
\end{aligned}$$

ϵ^2 is minimized by making $f(x) = v$, the orthogonal projection of x onto V. □

Dimensional Reduction by Orthogonal Projection

Corollary

Let $x_1, \ldots, x_K \in S$. Let V be a subspace of S. Let $f : S \to V$. Then

$$\min_f \sum_{k=1}^{K} (x_k - f(x_k))'(x_k - f(x_k))$$

is achieved when f is the orthogonal projection operator from S to V

Proof.

The best f can do for each x_k is for $f(x_k) = v_k$, the orthogonal projection of x_k onto V. Therefore,

$$\min_f \sum_{k=1}^{K} (x_k - f(x_k))'(x_k - f(x_k))$$

is achieved when f is the orthogonal projection operator onto V. □

Orthogonal Projection Operators

Proposition

If P is an orthogonal projection operator onto M dimensional subspace V of S, then for some orthonormal matrix T whose first M columns constitute an orthonormal basis for V,

$$P = TDT'$$

where D is a diagonal matrix whose first M diagonal entries are 1 and whose remaining diagonal entries are 0.

Proof.

Let $b_1, \ldots, b_N$ be an orthonormal basis of S with $b_1, \ldots, b_M$ being an orthonormal basis of V. Then $P = \sum_{m=1}^{M} b_m b_m'$. Let $T = (b_1 b_2 \ldots b_N)$ Consider,

$$TDT' = \begin{pmatrix} \vdots & \ldots & \vdots \\ b_1 & \ldots & b_N \\ \vdots & \ldots & \vdots \end{pmatrix} \begin{pmatrix} 1 & & & & & & \\ & 1 & & & & & \\ & & \vdots & & & & \\ & & & 1 & & & \\ & & & & 0 & & \\ & & & & & \vdots & \\ & & & & & & 0 \end{pmatrix} \begin{pmatrix} \ldots & b_1' & \ldots \\ \ldots & \vdots & \ldots \\ \ldots & b_N' & \ldots \end{pmatrix}$$

Orthogonal Projection Operators

$TDT' =$

$$\begin{pmatrix} \vdots & \ldots & \vdots & 0\ldots & 0 & \\ b_1 & \ldots & b_M & \vdots & \ldots & \vdots \\ \vdots & \ldots & \vdots & 0\ldots & 0 & \end{pmatrix} \begin{pmatrix} \ldots & b_1' & \ldots \\ \ldots & \vdots & \ldots \\ \ldots & b_N' & \ldots \end{pmatrix}$$

$$= \begin{pmatrix} \vdots & \vdots & \ldots & \vdots \\ b_1 & b_2 & \ldots & b_M \\ \vdots & \vdots & \ldots & \vdots \end{pmatrix} \begin{pmatrix} \ldots & b_1' & \ldots \\ \ldots & b_2' & \ldots \\ \vdots & \vdots & \vdots \\ \ldots & b_M' & \ldots \end{pmatrix}$$

$$= \sum_{m=1}^{M} b_m b_m'$$

Orthogonal Projection Operator Example

Consider the orthogonal projection operator onto the space spanned by

$$\frac{1}{5}\begin{pmatrix} 3 \\ 4 \end{pmatrix}$$

$$P = \frac{1}{5}\begin{pmatrix} 3 \\ 4 \end{pmatrix}\frac{1}{5}(3\ 4) = \frac{1}{25}\begin{pmatrix} 9 & 12 \\ 12 & 16 \end{pmatrix}$$

$$\frac{1}{25}\begin{pmatrix} 9 & 12 \\ 12 & 16 \end{pmatrix} = \begin{pmatrix} \frac{3}{5} & \frac{-4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{-4}{5} & \frac{3}{5} \end{pmatrix}$$

Orthogonal Projection Operators

Proposition

Let P be an orthogonal projection operator and T be an orthonormal matrix. Then $Q = TPT'$ is an orthogonal projection operator.

Proof.

$$
\begin{aligned}
Q^2 &= (TPT')(TPT') \\
&= TP(T'T)PT' \\
&= TP^2T' \\
&= TPT' = Q \\
Q' &= (TPT')' \\
&= TP'T' = TPT'
\end{aligned}
$$

□

Orthogonal Projection Operators

Proposition

Let P be an orthogonal projection operator. Then the diagonal elements of P lie in the interval $[0, 1]$

Proof.

Since $P^2 = P$, $p_{ij} = \sum_{n=1}^{N} p_{in}p_{nj}$. In particular $p_{ii} = \sum_{n=1}^{N} p_{in}p_{ni}$. Since $P = P'$, $p_{ii} = \sum_{n=1}^{N} p_{in}p_{in}$. Now, $p_{ii} = \sum_{n=1}^{N} p_{in}^2$ implies $p_{ii} \geq 0$. And $p_{ii} = p_{ii}^2 + \sum_{\substack{n=1 \\ n \neq i}}^{N} p_{in}^2$ implies $p_{ii} \geq p_{ii}^2$ from which $p_{ii} \leq 1$. □

Kernel and Range

Definition

- The Kernel of a matrix operator A is

$$Kernel(A) = \{x \mid Ax = 0\}$$

- The Range of a matrix operator A is

$$Range(A) = \{y \mid \text{for some } \; x, y = Ax\}$$

Kernel and Range

Proposition

Let P be a projection operator onto subspace V of S. Then

$$Range(P) + Ker(P) = S$$

Proof.

Let $x \in S$. $Px + (I - P)x = Px + x - Px = x$. Certainly $Px \in Range(P)$. Consider $(I - P)x$. $P[(I - P)x] = Px - PPx = Px - Px = 0$ Therefore, by definition of $Kernel(P)$, $(I - P)x \in Kernel(P)$. □

Kernel and Range

Proposition

Let P be an orthogonal projection operator. Then $Range(P) \perp Kernel(P)$

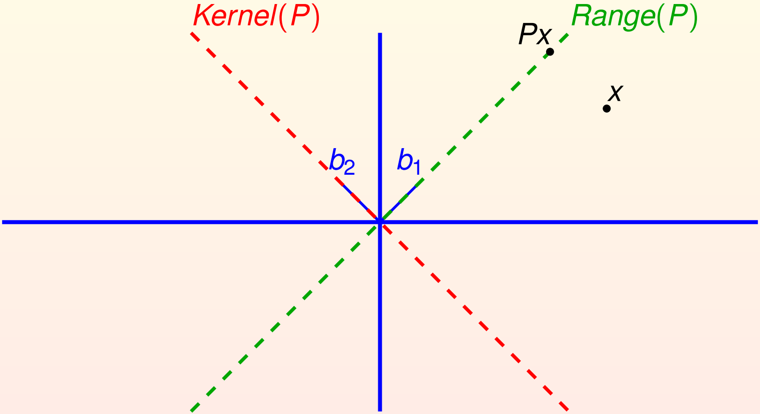
Proof.

Let $x \in Range(P)$ and $y \in Kernel(P)$. Then for some u, $x = Pu$. Consider $x'y$.

$$x'y = (Pu)'y = u'P'y = u'Py$$

But $y \in Kernel(P)$ so that Py=0. Therefore $x'y = 0$. □

Projecting



$$P = \begin{pmatrix} .5 & .5 \\ .5 & .5 \end{pmatrix}$$

Proposition

Let P be the orthogonal projection operator onto the subspace V. Then $I - P$ is the orthogonal projection operator onto the subspace $V^\perp$.

Proof.

$$\begin{aligned} (I-P)(I-P) &= I - P - P + P^2 = I - 2P + P = I - P \\ (I-P)' &= I' - P' = I - P \end{aligned}$$

$V^\perp = Kernel(P)$. Let $x \in V^\perp$. Then $Px = 0$. Consider $(I-P)x = x - Px = x$ □

Trace

Definition

Let $A = (a_{ij})$ be a square $N \times N$ matrix.

$$Trace(A) = \sum_{n=1}^{N} a_{nn}$$

Proposition

$Trace(\sum_{n=1}^{N} \alpha_n A_n) = \sum_{n=1}^{N} \alpha_n Trace(A_n)$

Trace

Proposition

$$Trace(AB) = Trace(BA)$$

Proof.

Let $C^{N\times N} = (c_{ij}) = A^{N\times K}B^{K\times N}$ and $D^{K\times K} = (d_{mn}) = B^{K\times N}A^{N\times K}$.

$$\begin{aligned}
c_{ij} &= \sum_{k=1}^{K} a_{ik}b_{kj} \\
d_{mn} &= \sum_{i=1}^{N} b_{mi}a_{in} \\
Trace(C) &= \sum_{i=1}^{N} c_{ii} = \sum_{i=1}^{N}\sum_{k=1}^{K} a_{ik}b_{ki} \\
&= \sum_{k=1}^{K}\sum_{i=1}^{N} b_{ki}a_{ik} = \sum_{k=1}^{K} d_{kk} = Trace(D) = Trace(BA)
\end{aligned}$$

□

Trace

Corollary

$$x'Ax = Trace(Axx')$$

Proof.

$$\begin{aligned} x'Ax &= Trace(x'Ax) = Trace(x'(Ax)) \\ &= Trace((Ax)x') = Trace(Axx') \end{aligned}$$

□

Trace

Proposition

Let $A = (a_{ij})$ be a $M \times N$ matrix. Then

$$\sum_{m=1}^{M} \sum_{n=1}^{N} a_{mn}^2 = Trace(AA')$$

Proof.

Let $B = (b_{ij}) = AA'$. Then $b_{ij} = \sum_{n=1}^{N} a_{in} a_{jn}$. Hence,
$b_{ii} = \sum_{n=1}^{N} a_{in} a_{in} = \sum_{n=1}^{N} a_{in}^2$. Therefore
$Trace(B) = Trace(AA') = \sum_{m=1}^{M} b_{mm} = \sum_{m=1}^{M} \sum_{n=1}^{N} a_{mn}^2$ □

Trace

Proposition

Let P be an orthogonal projection operator to the M dimensional subspace V. Then $Trace(P) = M$

Proof.

Let $b_1, \ldots, b_M$ be an orthonormal basis for V. Then $P = \sum_{m=1}^{M} b_m b'_m$

$$\begin{aligned} Trace(P) &= Trace(\sum_{m=1}^{M} b_m b'_m) \\ &= \sum_{m=1}^{M} Trace(b_m b'_m) = \sum_{m=1}^{M} Trace(b'_m b_m) \\ &= \sum_{m=1}^{M} Trace(1) = \sum_{m=1}^{M} 1 = M \end{aligned}$$

Trace

Proposition

Let P be an orthogonal projection operator onto a M dimensional subspace. Then

$$\sum_{i=1}^{N}\sum_{j=1}^{N} p_{ij}^2 = M$$

Proof.

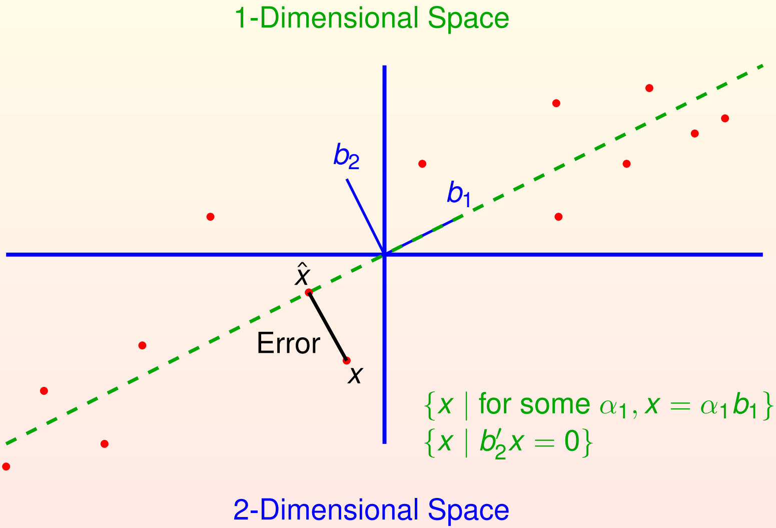
$$\sum_{i=1}^{N}\sum_{j=1}^{N} p_{ij}^2 = Trace(PP') = Trace(PP) = Trace(P) = M$$

□

Representing Subspaces

- When we squeeze N dimensional data into an M dimensional subspace, our viewpoint is the M degrees of freedom perspective
- When we fit N dimensional data to an $N - M$ dimensional function our viewpoint is the $N - M$ degrees of constraint perspective

Line Fitting



Curve Fitting

- $f(x) = 0$ are the constraints specifying the subspace
- f takes N Dimensional vectors to $N - M$ Dimensional subspaces in the N-Dimensional Space
- N dimensional vectors $x_1, \ldots, x_K$
- Fitting function f
- Squared Error $\epsilon^2 = \sum_{k=1}^{K} f(x_k)' f(x_k)$
- Find f to minimize ϵ^2

Dimensionality Reduction

- $f(x)$ is the dimensionality reduced vector
- f takes N Dimensional vectors to M Dimensional subspaces in the N-Dimensional Space
- N dimensional vectors $x_1, \ldots, x_K$
- Squared Error $\epsilon^2 = \sum_{k=1}^{K} (x_k - f(x_k))'(x_k - f(x_k))$
- Find f to minimize ϵ^2

Orthogonal Projection Operators are Positive Semidefinite

Proposition

Let $x \in S$ and P be an orthogonal projection operator. Then $x'Px \geq 0$.

Proof.

Since P is an orthogonal projection operator, $P = P^2$ and $P = P'$. Then,

$$x'Px = x'PPx = x'P'Px = (Px)'Px \geq 0$$

□

Principal Components

Proposition

Let $x_1, \ldots, x_K \in S$ an N-dimensional vector space. Let P be an orthogonal projection operator having rank M. Then P minimizes $\sum_{k=1}^{K}(x_k - Px_k)'(x_k - Px_k)$ if and only if P maximizes $\sum_{k=1}^{K} x_k' P x_k$

Proof.

$$
\begin{aligned}
\sum_{k=1}^{K}(x_k - Px_k)'(x_k - Px_k) &= \sum_{k=1}^{K}(x_k' x_k - x_k' P x_k - x_k' P' x_k + x_k' P' P x_k \\
&= \sum_{k=1}^{K}(x_k' x_k - x_k' P x_k - x_k' P x_k + x_k' P x_k \\
&= \sum_{k=1}^{K} x_k' x_k - \sum_{k=1}^{K} x_k' P x_k
\end{aligned}
$$

□

Principal Components

Proposition

Let $x_1, \ldots, x_K \in S$, an N-dimensional vector space and Q be an orthogonal projection operator of rank M. Then

$$\sum_{k=1}^{K} x_k Q x_k = Trace(Q^* D)$$

where TDT' is the eigenvector representation of $\sum_{k=1}^{K} x_k x_k'$ and $Q^* = T'QT$.

Proof.

$$\begin{aligned}\epsilon^2 = \sum_{k=1}^{K} x_k' Q x_k &= \sum_{k=1}^{K} Trace(x_k' Q x_k) = \sum_{k=1}^{K} Trace(Q x_k x_k') \\ &= Trace(\sum_{k=1}^{K} Q x_k x_k') = Trace(Q \sum_{k=1}^{K} x_k x_k')\end{aligned}$$

$\sum_{k=1}^{K} x_k x_k'$ is a real symmetric non-negative matrix. Therefore for some orthonormal matrix T and non-negative diagonal matrix D, $\sum_{k=1}^{K} x_k x_k' = TDT'$. Hence $\epsilon^2 = Trace(QTDT') = Trace((T'QT)D) = Trace(Q^* D)$ where $Q^* = T'QT$ □

Principal Components

Proposition

Let D be a diagonal matrix satisfying $d_{ii} \geq d_{jj}$, $i < j$. Let P be a rank M orthogonal projection operator. Then

$$\max_{\substack{P \\ P=P^2;P=P';Trace(P)=M}} Trace(PD) = \sum_{m=1}^{M} d_{mm}$$

Proof.

Consider $w_{ij} = \sum_{n=1}^{N} p_{in} d_{nj}$, the $(i,j)^{th}$ element of PD. Since $d_{nj} = 0$ for $n \neq j$, $w_{ij} = p_{ij} d_{jj}$. Hence $w_{ii} = p_{ii} d_{ii}$. Therefore $Trace(PD) = \sum_{i=1}^{N} p_{ii} d_{ii}$. Since P is an orthogonal projection operator, $0 \leq p_{ii} \leq 1$. $Trace(P) = M$ implies $\sum_{i=1}^{N} p_{ii} = M$. Since $d_{ii} \geq d_{jj}$, $i < j$, $\max_P Trace(PD) = \sum_{m=1}^{M} d_{mm}$. □

Principal Components

Theorem

Let $x_1, \ldots, x_K \in S$ an N-dimensional vector space and Q be an orthogonal projection operator of rank M. Then $\sum_{k=1}^{K} x_k Q x_k$ is maximized when Q projects to the M-Dimensional subspace spanned by the M eigenvectors of $\sum_{k=1}^{K} x_k x_k'$ having largest eigenvalues.

Proof.

Let $\sum_{k=1}^{K} x_k x_k' = TDT'$ and $Q^* = T'QT$. Without loss of generality we assume that the diagonal entries are ordered $d_{ii} \geq d_{jj}$, $i < j$. Then $\max_{Q^*} Trace(Q^* D) = \sum_{m=1}^{M} d_{mm}$, where the maximum is taken over all Q^* satisfying $Q^* = Q^* Q^*$ and $Q^* = Q^{*'}$. Thus, the first M diagonal entries of Q^* are one and the remaining diagonal entries 0. Since $\sum_{i=1}^{N} \sum_{j=1}^{N} q_{ij}^2 = M$, and there are M ones on the diagonal, the remaining elements of Q^* are 0. This implies $Q = TQ^* T'$ is the orthogonal projection operator onto the space spanned by the first M eigenvectors of $\sum_{k=1}^{K} x_k x_k'$ for these are the eigenvectors having largest eigenvalues. □