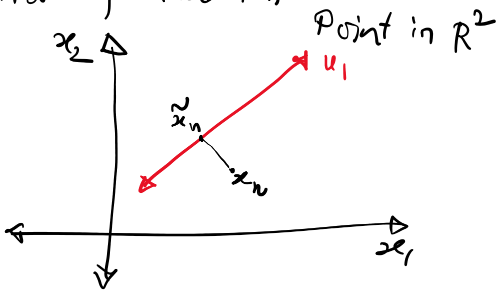
Minimum-error formulation

$\{x_n\}_1^N$:



Problem: Find a new set of orthogonal basis in M dimensional subspace (< D), such that projection error is minimized

* Consider a set of new ~~basis~~ vectors $\{u_1, u_2, \ldots, u_D\}$ (complete basis)
  Orthonormal basis

  $\therefore$ $u_i^T u_j = \delta_{ij} \rightarrow \delta_{ij} = \begin{cases} 1, & i=j \\ 0, & i \neq j \end{cases}$

* $x_n = \alpha_{n_1} u_1 + \alpha_{n_2} u_2 + \ldots + \alpha_{n_D} u_D$

  co-efficients for different points

* Compute $x_n^T u_j = \alpha_{n_1} \underbrace{u_1^T u_j}_{0} + \alpha_{n_2} \underbrace{u_2^T u_j}_{0} + \alpha_{n_j} \underbrace{u_j^T u_j}_{1} + \ldots + \alpha_{n_D} \underbrace{u_D^T u_j}_{0}$

  $= \alpha_{nj}$

* For the j^{th} α, $\alpha_{nj} = x_n^T u_j = u_j^T x_n$

* $x_n = \sum_{i=1}^{D} \alpha_{n_i} u_i = \sum_{i=1}^{D} (x_n^T u_i) u_i$

* Our Interest, finding $M (< D)$ variables

  $\tilde{x}_n = \sum_{i=1}^{M} \underbrace{z_{n_i}}_{\text{change with the point}} u_i + \sum_{i=M+1}^{D} \underbrace{b_i}_{\text{constant for a point}} u_i$

$\therefore$ Then, minimize error between x_n and $\tilde{x}_n$

N

∴ Then, minimize error between x_n and $\tilde{x}_n$

$$\min_{z_{nj}, b_j, u_j} J = \frac{1}{N} \sum_{n=1}^{N} \| x_n - \tilde{x}_n \|_2^2$$

$$= \frac{1}{N} \sum_{n=1}^{N} \| x_n - \sum_{i=1}^{M} z_{ni} u_i - \sum_{i=M+1}^{D} b_i u_i \|_2^2$$

* Determine z_{nj} and b_j by computing

$$\frac{\partial J}{\partial z_{nj}} = 0 \quad \text{and} \quad \frac{\partial J}{\partial b_j} = 0$$

→ $\frac{\partial J}{\partial z_{nj}} = 0 \iff \frac{\partial \tilde{x}_n}{\partial z_{nj}}^T (x_n - \tilde{x}_n) = 0$

$$\iff u_j^T \left(x_n - \sum_{i=1}^{M} z_{ni} u_i - \sum_{i=M+1}^{D} b_i u_i \right) = 0$$

$$\iff u_j^T x_n - z_{nj} = 0$$

$$\iff z_{nj} = u_j^T x_n \text{ or } x_n^T u_j, \quad j = 1, \dots, M$$

→ $\frac{\partial J}{\partial b_j} = 0 \iff \frac{\partial \tilde{x}_n}{\partial b_j}^T (x_n - \tilde{x}_n) = 0$

$$\iff b_j = \bar{x}_n^T u_j, \quad j = M+1, \dots, D$$

Page 2

12 August 2023 17:31

$$x_n - \tilde{x}_n = x_n - \sum_{i=1}^{M} z_{n_i} u_i - \sum_{i=M+1}^{D} b_i u_i$$

$$= \sum_{i=1}^{D} \alpha_{n_i} u_i - \sum_{i=1}^{M} z_{n_i} u_i - \sum_{i=1}^{D} b_i u_i$$

$$= \cancel{\sum_{i=1}^{M} (x_n^T u_i) u_i} + \sum_{i=M+1}^{D} (x_n^T u_i) u_i - \cancel{\sum_{i=1}^{M} (x_n^T u_i) u_i}$$

$$- \sum_{i=M+1}^{D} (\bar{x}_n^T u_i) u_i$$

$$= \sum_{i=M+1}^{D} \left\{ (x_n^T - \bar{x}_n^T) u_i \right\} u_i$$

* Let $\|x_n - \tilde{x}_n\|_2^2 = (x_n - \tilde{x}_n)^T (x_n - \tilde{x}_n)$

$$= \sum_{i=M+1}^{D} \left\{ u_i^T (x_n - \bar{x}_n) \right\} \overset{1}{u_i^T u_i} \left\{ (x_n - \bar{x}_n)^T u_i \right\}$$

$$= \sum_{i=M+1}^{D} u_i^T (x_n - \bar{x}_n)(x_n - \bar{x}_n)^T u_i$$

* $$J = \frac{1}{N} \sum_{n=1}^{N} \sum_{i=M+1}^{D} u_i^T (x_n - \bar{x}_n)(x_n - \bar{x}_n)^T u_i$$

$$= \sum_{i=M+1}^{D} u_i^T \left\{ \underbrace{\frac{1}{N} \sum_{n=1}^{N} (x_n - \bar{x}_n)(x_n - \bar{x}_n)^T}_{S} \right\} u_i$$

$$J = \sum_{i=1}^{D} u_i^T S u_i$$

* $$\boxed{\underset{u_i}{\text{minimize}}\ J = \sum_{i=M+1}^{D} u_i^T S u_i} \quad \text{— I}$$

* Let us take $D = 2$ and $M = 1$ for a simplification
* The objective function

$$\underset{u_2}{\text{minimize}}\ J = u_2^T S u_2$$

$$\text{s.t} \quad u_2^T u_2 = 1$$

{ otherwise trivial answer $\|u_2\| = 0$ }

* Lagrangian

$$\min_{u_2} \quad u_2^T S u_2 + \lambda (1 - u_2^T u_2)$$

$$\min_{u_2} \; u_2^T S u_2 + \lambda (1 - u_2^T u_2)$$

* $\frac{\partial J}{\partial u_2} = 0 \iff \boxed{S u_2 = \lambda_2 u_2}$.

* $J = u_2^T S u_2 + \lambda_2 (1 - u_2^T u_2)$

$= u_2^T \lambda_2 u_2 = \lambda_2$

* J takes a minimum value at $\lambda_2 \rightarrow$ by choosing smaller eigenvectors and

Then, to achieve the minimum error projection, in $D = 2$, and $M = 1$, the Principal component subspace corresponding to the ~~largest~~ eigenvalue should be chosen.

* For D-dimensional vector; and $M < D$.
Then, by taking $\{u_i\}$, $i = 1, \dots M$ corresponding to the M largest eigenvalues, the error objective fn

$$J = \sum_{i=M+1}^{D} \lambda_i \text{ is minimum.}$$

* PCs $\rightarrow$ the M eigenvectors corresponding to the largest eigenvalues

* orthogonal space: $(D - M)$ eigenvectors corresponding to the smallest Eigenvalues.