# Principal Component Analysis

The principal components of a matrix $X = U \Sigma V^T$ are in the columns of $V$.

By including a feature map, we can extend PCA to non linear spaces.

Assuming the data matrix is centered $X = \begin{bmatrix} \tilde{x}_1 \\ \vdots \\ \tilde{x}_n \end{bmatrix}$ we can write it in its singular value basis

$$Z = \begin{pmatrix} \lambda_1^T \\ \vdots \\ \lambda_n^T \end{pmatrix} = \begin{pmatrix} V^T \tilde{x}_1 \\ \vdots \\ V^T \tilde{x}_n \end{pmatrix} = \tilde{X} V$$

$$= \tilde{X} V D U^T U D^{-1} = \tilde{X} \tilde{X}^T U D^{-1}$$

$$= \tilde{K} U D^{-1}$$

kernel matrix for its feature space.

So the coordinates of the values in the new space can be rewritten as $\lambda_{im} = \sum K_{ij} u_{jm} d_m^{-1}$

With a feature map $h: \mathbb{R}^p \mapsto \mathbb{R}^m$ we can think of having a new training set $\tilde{X} = \begin{bmatrix} h(x_1) \\ \vdots \\ h(x_n) \end{bmatrix}$ which gives rise to another Principle Space.

and the coordinates in this space are given by

$$\lambda = \sum_{j=1} u_{jm} d_m^{-1} K(\tilde{x}, \tilde{x}_j)$$

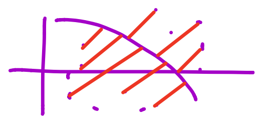
## Non linear PCA

Replace $x = f(\lambda)$ for some function

Suppose we have some distribution on $X \sim P_X(x)$

Define the projection

$$P(x) = \min_{\lambda} \|x - f(\lambda)\|$$



## Consistency Condition:

$$f(\lambda) = \mathbb{E}\left[X \mid P(x) = f(\lambda)\right]$$

Goal: find $f$ wrt this condition.

Inspires an iteration method

1. $\hat{f}(\lambda) = \mathbb{E}[X \mid \hat{\lambda}_f(x) = \lambda]$

2. $\hat{\lambda}_f(x) = \underset{\lambda}{\arg\min} \|x - f(\lambda)\|$

when $\hat{\lambda}_f(x) = \lambda$ s.t. $f(\lambda) = P(x)$

## Graph Clustering Methods

Define a similarity metric $d(x_i, x_j) = \exp\left\{-\|x_i - x_j\|^2 / c\right\}$

Prune tree to only include 5 largest $w_{ij} = d(x_i, x_j)$

Then define the Graph Laplacian $L = D - W$