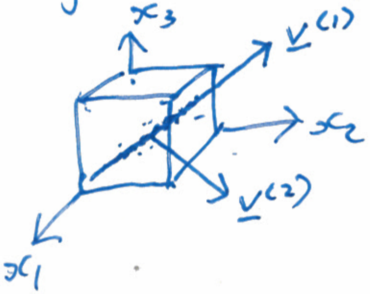
PCA

High-dim ball of points



$\underline{v}^{(k)}$ eigenvector of $\mathrm{cov}[X]$

$$\underset{D\times K}{V} = \begin{bmatrix} | & | & & | \\ \underline{v}^{(1)} & \underline{v}^{(2)} & \cdots & \underline{v}^{(k)} \\ | & | & & | \end{bmatrix}$$

Individual vector

$$\underline{x} \longrightarrow \underbrace{V^T \underline{x}}_{K\times 1}$$

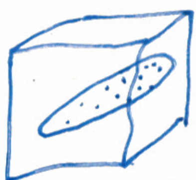
$N\times D$ Design matrix:

$$X \longrightarrow \underbrace{\underset{N\times D}{X}\ \underset{D\times K}{V}}_{N\times K}$$

Lossily reconstruct

$$\hat{X} = \underbrace{X V V^T}_{N\times D}$$

If $K=2$



All reconstructed points exactly on K-dim subspace.

Gaussian model for PCA

Assume there is a process in k-dim:

$$\underline{h}^{(n)} \sim N(0, I_K)$$

$$\underline{x}^{(n)} = \underbrace{W}_{D \times K} \underline{h}^{(n)} + \text{Gaussian noise, zero mean, covariance } \sigma_n^2 I$$

These lie exactly on K-dim linear subspace

$$\underline{x} \sim N(0, WW^T + \sigma_n^2 I)$$

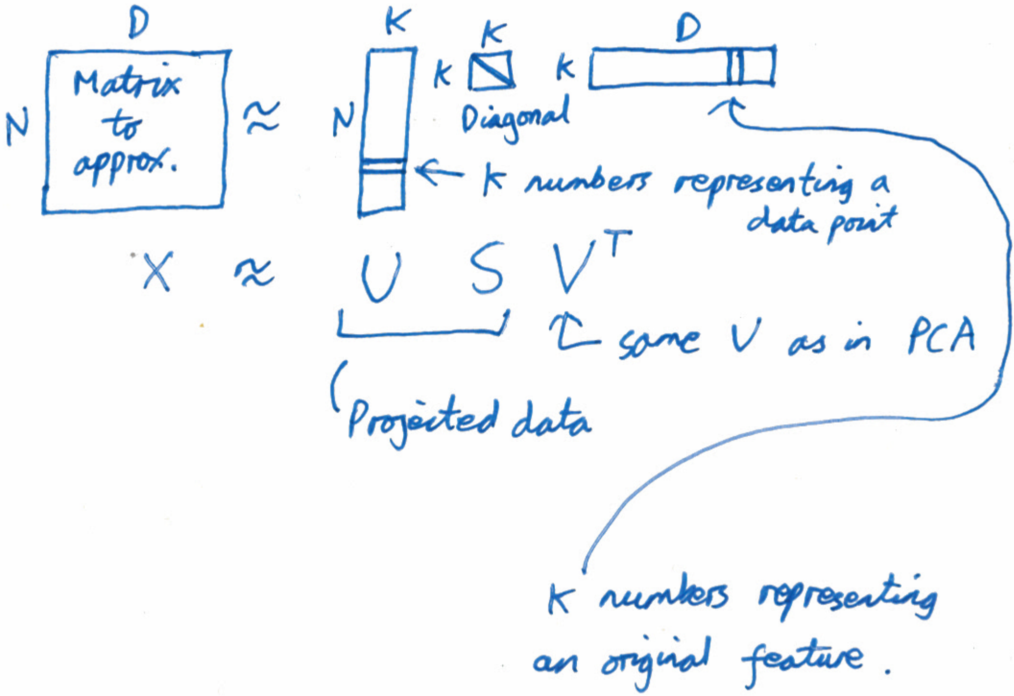
$$\begin{aligned} &\mathrm{cov}[Wh^{(n)}] \\ &= \mathbb{E}[W\underline{h}^{(n)}(W\underline{h}^{(n)})^T] \\ &= W \underbrace{\mathbb{E}[\underline{h}^{(n)}\underline{h}^{(n)T}]}_{I} W^T \\ &= WW^T \end{aligned}$$

"Probabalistic PCA"

"Factor Analysis" where noise has arbitrary diagonal covariance

Truncated SVD

SVD is some standard linear algebra method



[See typeset notes for links to papers with figures.]

[Also for discussion of privacy.]

Truncated SVD

```
% PCA via SVD,
% for zero-mean X:
[U, S, V] = svd(X, 0);
U = U(:, 1:K);
S = S(1:K, 1:K);
V = V(:, 1:K);
X_kdim = U*S;
X_proj = U*S*V';
```

$$\begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1D} \\ X_{21} & X_{22} & \cdots & X_{2D} \\ X_{31} & X_{32} & \cdots & X_{3D} \\ X_{41} & \boxed{X_{42}} & \cdots & X_{4D} \\ X_{51} & X_{52} & \cdots & X_{5D} \\ \vdots & \vdots & \ddots & \vdots \\ X_{N1} & X_{N2} & \cdots & X_{ND} \end{bmatrix} \approx$$

$$\begin{bmatrix} U_{11} & \cdots & U_{1K} \\ U_{21} & \cdots & U_{2K} \\ U_{31} & \cdots & U_{3K} \\ \boxed{U_{41}} & \boxed{\cdots} & \boxed{U_{4K}} \\ U_{51} & \cdots & U_{5K} \\ \vdots & \ddots & \vdots \\ U_{N1} & \cdots & U_{NK} \end{bmatrix} \begin{bmatrix} \boxed{S_{11}} & 0 & 0 \\ 0 & \boxed{\ddots} & 0 \\ 0 & 0 & \boxed{S_{KK}} \end{bmatrix} \begin{bmatrix} V_{11} & \boxed{V_{21}} & \cdots & V_{D1} \\ \vdots & \boxed{\vdots} & \ddots & \vdots \\ V_{1K} & \boxed{V_{2K}} & \cdots & V_{DK} \end{bmatrix}$$

$$X \approx U \quad S \quad V^\top$$