Applied Machine Learning

Principal Component Analysis

Principal Component Analysis

- Translation and rotations
- Dimensionality reduction
- Error in low-dimensional representation of high-dimensional data

Data Transformation

- Source dataset: $\{\mathbf{x}\}$
- After translation: $\{\mathbf{m}\}$
  - $\mathbf{m}_i = \mathbf{x}_i - mean(\{\mathbf{x}\})$
- And, after rotation: $\{\mathbf{r}\}$
  - Covmat$\{\mathbf{r}\} = U^\top \Sigma U = \Lambda$
  - $\mathbf{r}_i = U^\top \mathbf{m}_i = U^\top(\mathbf{x}_i - \text{mean}(\{x\}))$
  - $mean(\{\mathbf{r}\}) = mean(\{\mathbf{m}\}) = 0$

- Eigenvectors U and eigenvalues Λ of S
  - $SU = U\Lambda$
- Diagonalization
  - $U^\top S U = \Lambda$
- $A = U^\top$
- Covmat$\{\mathbf{x}\} = \Sigma = S$

Selection of Features

- Eigenvalues in Λ
  - sorted from largest to smallest
  - the larger the component, the larger the weight for the corresponding feature
- New dataset $\{\mathbf{p}\}$
  - Select s largest components in Λ and corresponding eigenvectors in U
  - Replace remaining $d - s$ components with 0

Error in New Dataset

- $\frac{1}{N}\sum_i \left[(\mathbf{r}_i - \mathbf{p}_i)^\top(\mathbf{r}_i - \mathbf{p}_i)\right]$
- $= \quad \frac{1}{N}\sum_i \left[\sum_{j=s+1}^{j=d} (r_i^{(j)})^2\right]$

  $= \quad \sum_{j=s+1}^{j=d} \left[\frac{1}{N}\sum_i (r_i^{(j)})^2\right]$

  $= \quad \sum_{j=s+1}^{j=d} \text{var}(\{r^{(j)}\})$
- $= \quad \sum_{j=s+1}^{j=d} \lambda_j$

- Relative error
  - $\frac{\sum_{j=s+1}^{j=d} \lambda_j}{\sum_{j=1}^{j=d} \lambda_j}$ should be small

Inverse Transformation

- Lower dimensional dataset $\{\hat{\mathbf{x}}\}$

$$\hat{\mathbf{x}}_i = U\mathbf{p}_i + mean(\{\mathbf{x}\})$$

- $= \sum_{j=1}^{j=s} r_i^{(j)}\mathbf{u}_j + mean(\{\mathbf{x}\})$

  - $\mathbf{r}_i == U^\top(\mathbf{x}_i - \text{mean}(\{x\}))$
  - $r_i^{(j)} = \mathbf{u}_j^\top(\mathbf{x}_i - \text{mean}(\{x\}))$

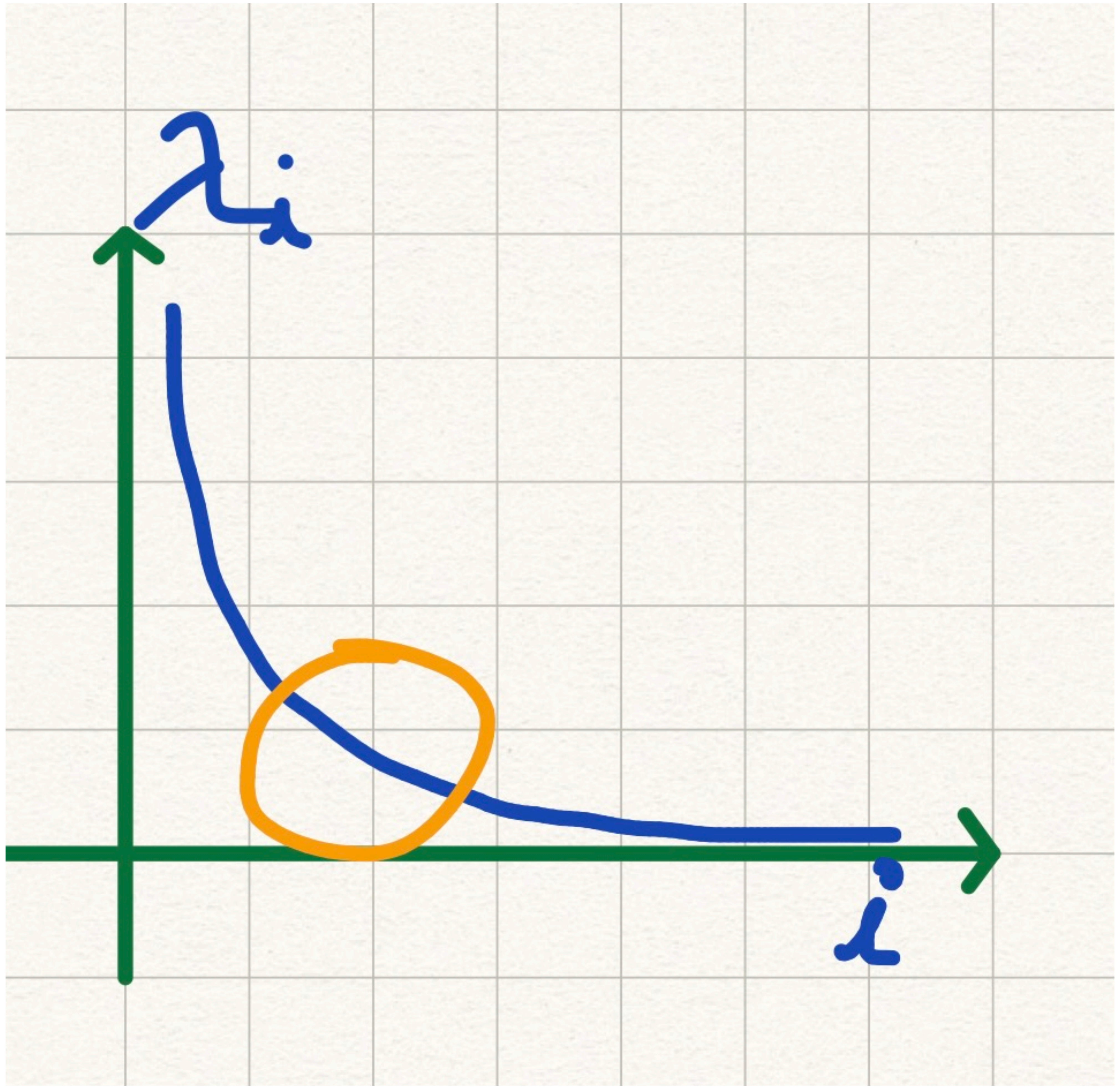
- $\hat{\mathbf{x}}_i = \sum_{j=1}^{j=s}\left[\mathbf{u}_j^\top(\mathbf{x}_i - \text{mean}(\{x\}))\mathbf{u}_j\right] + mean(\{\mathbf{x}\})$

- Mean error

- $\displaystyle\sum_{j=s+1}^{j=d} \lambda_j$

Principal Component Analysis

- Original dataset: $\{\mathbf{x}\}$
  - d features
  - $U^\top \text{Covmat}\{\mathbf{x}\} U = U^\top \Sigma U = \Lambda$
- Choose s features
  - small ratio in $\dfrac{\sum_{j=s+1}^{j=d} \lambda_j}{\sum_{j=1}^{j=d} \lambda_j}$
  - plot relative error vs s or λ_i vs i, select s most significant
- Low-Dimensional representation $\hat{\mathbf{x}}$:
  - $\hat{\mathbf{x}}_i \quad = \quad \sum_{j=1}^{j=s} \left[\mathbf{u}_j^\top (\mathbf{x}_i - \text{mean}(\{x\}))\mathbf{u}_j \right] + mean(\{\mathbf{x}\})$



Principal Component Analysis

- Translation and rotations
- Dimensionality reduction
- Error in low-dimensional representation of high-dimensional data

Applied Machine Learning

Principal Component Analysis