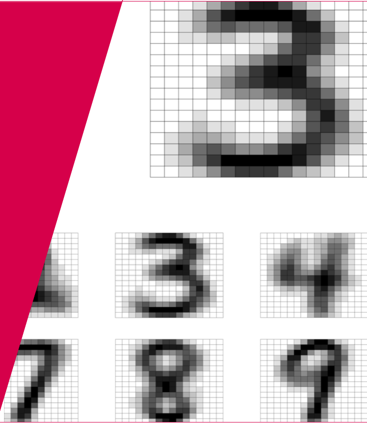
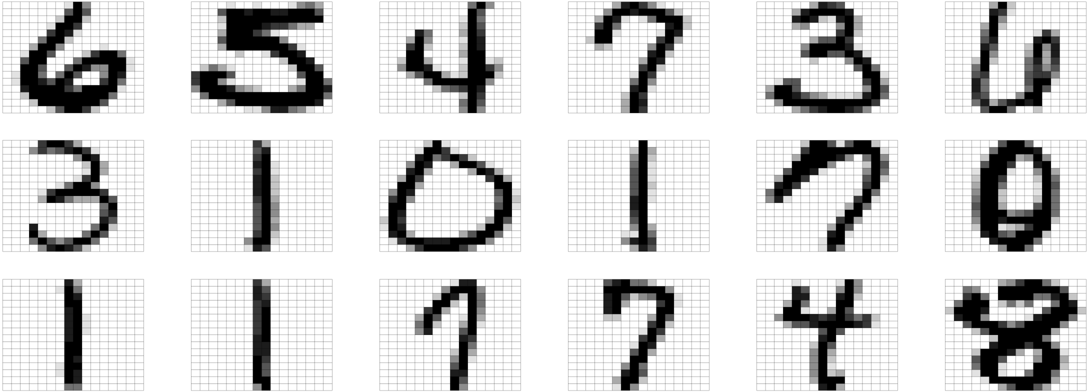
Recognizing digits written by hand

Martijn Anthonissen



Technische Universiteit
Eindhoven
University of Technology

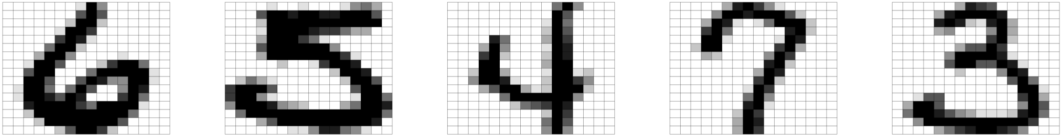
Maths can teach a computer to read zipcodes



Technische Universiteit Eindhoven University of Technology

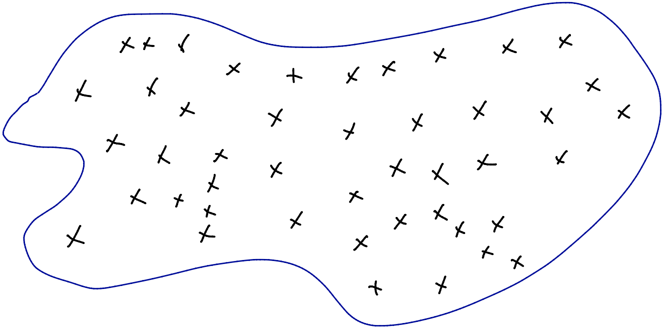
Problem statement

- **Given**: 1707 training images
  Each image is 16×16 gray scale image
  Each image has been *manually classified*: We know which digit it is
- **Goal**: Classify new unknown images by computer



Mathematical model

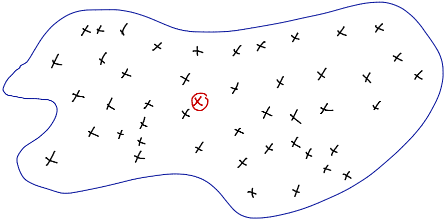
- Each image has $16 \times 16 = 256$ pixels $\Longrightarrow$ Vector $\mathbf{a} \in \mathbb{R}^{256}$



- Distance between two images:

$$d(\mathbf{a}, \mathbf{b}) = \|\mathbf{a} - \mathbf{b}\|_2 = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \cdots + (a_{256} - b_{256})^2}$$

Classification based on nearest known image

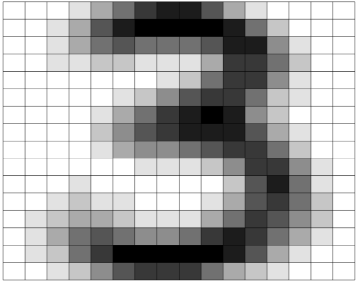


Black x: Known image
Red circled x: Unknown image

Method 1: Nearest known image

1. Compute distances of unknown image to all 1707 known images
2. Find nearest known image
3. Classify unknown image as digit in known image

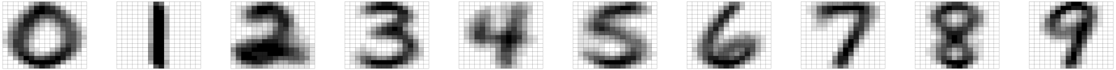
Compute the mean image for all digits



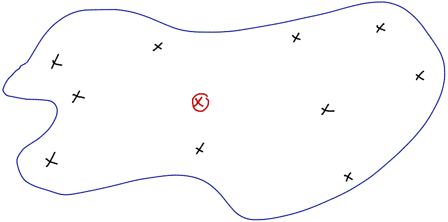
Start

- Collect all 3’s in known images
- Compute the mean

Same for other digits



Classification based on nearest mean



Black x: Mean image
Red circled x: Unknown image

Method 2: Nearest mean

1. Compute distances of unknown image to the 10 mean images
2. Find nearest mean image
3. Classify unknown image as digit in nearest mean

Recap singular value decomposition

$$\mathbf{A} = \mathbf{U} \, \boldsymbol{\Sigma} \, \mathbf{V}^T$$

$\mathbf{A}$: $m \times n$

$\mathbf{U}$: $m \times m$, orthogonal

columns of $\mathbf{U}$ are the eigenvectors of $\mathbf{A}\mathbf{A}^T$

columns of $\mathbf{U}$ are the *left singular vectors* of $\mathbf{A}$

$\boldsymbol{\Sigma}$: $m \times n$, diagonal

entries on diagonal are the *singular values* of $\mathbf{A}$

$\mathbf{V}^T$: $n \times n$, orthogonal

columns of $\mathbf{V}$ are the eigenvectors of $\mathbf{A}^T\mathbf{A}$

columns of $\mathbf{V}$ are the *right singular vectors* of $\mathbf{A}$

Application of the SVD: Approximating a matrix

$$\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^T = (\mathbf{u}_1 \mid \cdots \mid \mathbf{u}_m) \begin{pmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_n \\ 0 & \cdots & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & \cdots & 0 \end{pmatrix} \begin{pmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \vdots \\ \mathbf{v}_n^T \end{pmatrix} = (\mathbf{u}_1 \mid \cdots \mid \mathbf{u}_m) \begin{pmatrix} \sigma_1 \mathbf{v}_1^T \\ \sigma_2 \mathbf{v}_2^T \\ \vdots \\ \sigma_n \mathbf{v}_n^T \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$= \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T + \cdots + \sigma_n \mathbf{u}_n \mathbf{v}_n^T$$

Theorem (Decomposition in rank-1 matrices)

Any matrix $\mathbf{A}$ can be decomposed as a sum of rank-1 matrices:

$$\mathbf{A} = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T + \cdots + \sigma_n \mathbf{u}_n \mathbf{v}_n^T$$

We have $\|\mathbf{u}_i \mathbf{v}_i^T\|_2 = 1$

Approximating a matrix

Theorem (Decomposition in rank-1 matrices)

Any matrix $\mathbf{A}$ can be decomposed as a sum of rank-1 matrices:

$$\mathbf{A} = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T + \cdots + \sigma_n \mathbf{u}_n \mathbf{v}_n^T$$

We have $\|\mathbf{u}_i \mathbf{v}_i^T\|_2 = 1$

Can be used to approximate matrices:
Define $\mathbf{\Sigma}_r = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_r, 0, 0, \ldots, 0)$ for $r \leq n$
The matrix

$$\mathbf{A}_r := \mathbf{U}\mathbf{\Sigma}_r\mathbf{V}^T = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T + \cdots + \sigma_r \mathbf{u}_r \mathbf{v}_r^T$$

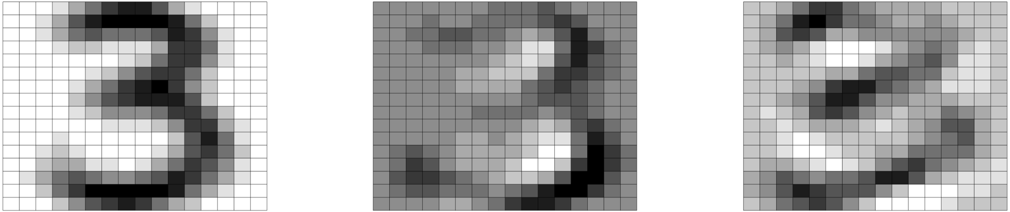
is a rank-r approximation of $\mathbf{A}$

Using the singular value decomposition

Training

- Collect all 3's in the training set. There are 131 images manually classified as 3
- Build a matrix $\mathbf{A} \in \mathbb{R}^{256 \times 131}$
- Compute the singular value decomposition $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$
- Select the first r columns of $\mathbf{U}$. Define $\mathbf{U}_r := (\mathbf{u}_1 \mid \mathbf{u}_2 \mid \cdots \mid \mathbf{u}_r)$

  First three columns:



- Same for the other nine digits

Approximation with the left singular vectors

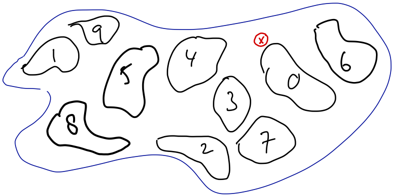
- Let $\mathbf{z}$ be an unknown image
- Approximate $\mathbf{z}$ using linear combination of $\mathbf{u}_1, \ldots, \mathbf{u}_r$: $\quad \mathbf{z} \approx \sum_{i=1}^{r} y_i \mathbf{u}_i = \mathbf{U}_r \mathbf{y} \qquad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_r \end{pmatrix}$
- Take inner product with $\mathbf{u}_j$ to find:

$$(\mathbf{z}, \mathbf{u}_j) \approx \sum_{i=1}^{r} y_i (\mathbf{u}_i, \mathbf{u}_j) = \sum_{i=1}^{r} y_i \delta_{ij} = y_j \qquad \mathbf{y} = \mathbf{U}_r^T \mathbf{z}$$

- Conclusion: $\mathbf{z} \approx \mathbf{U}_r \mathbf{U}_r^T \mathbf{z}$ This is the orthogonal projection of $\mathbf{z}$ on $\langle \mathbf{u}_1, \ldots, \mathbf{u}_r \rangle$
- Define the relative residual as: $\dfrac{\|\mathbf{z} - \mathbf{U}_r \mathbf{U}_r^T \mathbf{z}\|_2}{\|\mathbf{z}\|_2}$

  Measures how well $\mathbf{z}$ can be approximated with the left singular vectors

Classification based on SVD



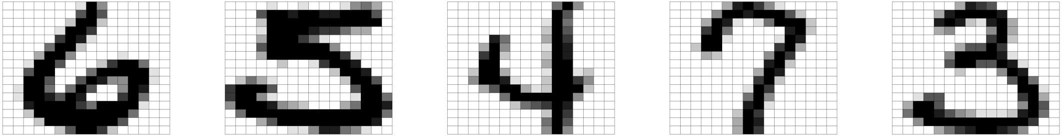
Black cloud: Combinations of left singular values of digit
Red circled x: Unknown image

Based on SVD

1. Compute first r left singular vectors for 0, 1, …, 9
2. Compute relative residual of unknown image expressed in all ten bases
3. Find minimum relative residual and classify unknown image as corresponding digit

Conclusions and further reading

- Images can be represented by vectors
- The SVD finds main features in given data
- Maths can teach a computer to read zipcodes!



Background material:

- Based on Chapter 10: Classification of Handwritten Digits from the book *Matrix Methods in Data Mining and Pattern Recognition* by Lars Eldén, SIAM, Philadelphia, 2007
- See http://users.mai.liu.se/larel04/matrix-methods/computer-assignments/character-recogn.html
- The data are a subset of the US Postal Service Database