Group discussion 1: PCA continued

Ex. 1 — (4 min) In the lecture, we discussed the PCA for MNIST dataset. Each image consists of 28×28 pixels and can be represented as a 784-dimensional vector. We analyze all the ”1”-images.

1. What is dimension of the mean of $\boldsymbol{x}$, i.e. $\widehat{\mu_x}$?

784

2. What is dimension of the covariance matrix C_x?

784*784 (features * features)

3. What is the dimension of the eigenvectors of C_x?

784

4. How many eigenvalues does C_x have?

784

Ex. 2 — (4 min) What is the total variance and the explained variance?
The total variance is the sum of all the eigenvalues in the covariance matrix:

$$T = \sum_i \lambda_i$$

Explained variance is the sum of the first n eigenvalues in the covariance matrix div. total variance.

$$Explained\ variance = \frac{\sum_{i=1}^{10} \lambda_i}{T}$$

Ex. 3 — (6 min) Argue with the following statements (are they correct or not, and why?):

1. We can represent any ”1”-image in the eigenbasis without introducing errors
Correct, the picture will merely the rotated. The eigenbasis is just another basis.

2. We can represent any ”1”-image in a subset of eigenbasis, but it’s only an approximation
Correct, this is the principle of PCA.

3. We can represent any digit in the eigenbasis from obtained ”1”-images
Correct, the eigenbasis is still a basis.

4. We can represent other digits in a subset of eigenbasis from obtained ”1”-images, and it will give a good approximation

It will not always be correct, but it may be in some cases.

Ex. 4 — (6 min) The eigenvectors of the matrix

$\boldsymbol{A}$ is $\boldsymbol{v}_1 = (1/\sqrt{2}, 1/\sqrt{2})^T$ and $\boldsymbol{v}_2 = (-1/\sqrt{2}, 1/\sqrt{2})^T$ is with corresponding eigenvales $\lambda_1 = 2$ and $\lambda_2 = 1$, respectively.

1. Which of the following are true or false?

$$
\begin{aligned}
\text{(I)}\ \boldsymbol{A}\boldsymbol{v}_2 &= (-1/\sqrt{2}, 1/\sqrt{2})^T \\
\text{(II)}\ \boldsymbol{A}\boldsymbol{v}_2 &= (1/\sqrt{2}, -1/\sqrt{2})^T \\
\text{(III)}\ \boldsymbol{A}\boldsymbol{v}_2 &= (-2/\sqrt{2}, 2/\sqrt{2})^T \\
\text{(IV)}\ \boldsymbol{A}\boldsymbol{v}_1 &= (-2/\sqrt{2}, 2/\sqrt{2})^T \\
\text{(V)}\ \boldsymbol{A}\boldsymbol{v}_1 &= (1/\sqrt{2}, 1/\sqrt{2})^T \\
\text{(VI)}\ \boldsymbol{A}\boldsymbol{v}_1 &= (2/\sqrt{2}, 2/\sqrt{2})^T
\end{aligned}
$$

$Av_1 = \lambda_1 v_1$

$Av_2 = \lambda_2 v_2$

(I) True

(II) False

(III) False

(IV) False

(V) False

(VI) True

2. Let $w = (1, 3)^T$. How can we represent w in terms of the eigenvectors? i.e. how can we express w in the basis $V = (v_1, v_2)$?

$w = 2\sqrt{2} * v_1 + \sqrt{2} * v_2$

Hint: Try to write $w = av_1 + bv_w$, where a and b are constants.

https://prod.liveshare.vsengsaas.visualstudio.com/join?54EE69670CF08D2F42AFF51391256ACADA11

Group discussion 2: Image, filters and convolutions

Ex. 5 — (3 min) We learned in lecture 1 that greyscale images can be represented by a N × M matrix. RGB images are color images, how are these represented? How many dimensions does an N ×M RGB image live in?
3 grayscale layers / channels of red, green and blue respectively.
N × M × 3

Ex. 6 — (7 min) We define the convolution as :

$$f_{\text{con-out}}(x,y) = w * f_{in}(x,y) = \sum_{v=-b}^{b}\sum_{u=-a}^{a} w(u,v) f_{in}(x-u, y-v)$$

and the cross-correlation as:

$$f_{\text{cro-out}}(x,y) = w * f_{in}(x,y) = \sum_{v=-b}^{b}\sum_{u=-a}^{a} w(u,v) f_{in}(x+u, y+v)$$

1. Show that output image from the cross correlation with the flipped kernel $w^f(-u,-v)$ equal to the output image from the convolution with the kernel itself.

$$f_{cro-out}(x,y) = \sum_{v=-b}^{b}\sum_{v=-a}^{a} w(u,v)f_{in}(x+u,\ y+v)$$

$$f_{cro-out-flip}(x,y) = w^f * f_{in} = \sum_{v=-b}^{b}\sum_{v=-a}^{a} w^f(u,v)f_{in}(x+u,\ y+v)$$

$$= \sum_{v=-b}^{b}\sum_{v=-a}^{a} w(-u,-v)f_{in}(x+u,\ y+v)$$

Changing variables
$u =- s,\ v =- r$

$$= \sum_{s=-b}^{b}\sum_{s=-a}^{a} w(s,\ v)f_{in}(x-s,\ y-v) = f_{con-out}(x,y)$$

2. How can one determine the value of a and b in the previous equations?

$$a = \frac{(w_{kernel}-1)}{2}$$

$$b = \frac{(h_{kernel}-1)}{2}$$

Ex. 7 — (5min) Going back to gray-scale images, explain in words how one can apply convolution in practice step by step.
Create a kernel

Elementwise multiplying the kernel value to the underlying input value, then summing all those products into one value, essentially combining all the information into a single value. Then combining each of the values found into a combined output, which is what is called a convolution. The convolution itself then describes/includes the most significant features in the input given the kernel.

Ex. 8 — (10 min) Calculate the following:

1. The valid output image f_{out} of the following convolution

$$f_{out} = w * f_{in}$$

with

$$w = \begin{bmatrix} 1 & 2 & 1 \end{bmatrix}$$
$$f_{in} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 3 & 0 \end{bmatrix}.$$

1	2	1

0	1	0
0	2	0
0	3	0

2
4
6

Padding is a way to get the output image the same size as the input image. There is a relation between the output image size with the input image size, the kernel size, number of zero padding (P) and number of stride* (S):

$$\text{output width} = \frac{\text{input width} - \text{kernel width} + 2P}{S} + 1$$

$$\text{output height} = \frac{\text{input height} - \text{kernel height} + 2P}{S} + 1$$

*Stride is the number of pixels shifts over the input matrix. When the stride is 1 then we move the filters to 1 pixel at a time. When the stride is 2 then we move the filters to 2 pixels at a time and so on.

2. Calculate the zero-padded output image f_{out} with the same kernel and input as above

0	0	1	0	0
0	0	2	0	0
0	0	3	0	0

$$\downarrow$$

1	2	1
2	4	2
3	6	3