TU München
Fakultät für Informatik
PD Dr. Rudolph Triebel
John Chiotellis

Machine Learning for Computer Vision
Summer term 2017

16. Mai 2017
Topic: Linear Algebra

Exercise 1: Warm up

a) What multiple of $a = (1,1,1)$ is closest to the point $b = (2,4,4)$? Find also the closest point to a on the line through b.

There is some vector $p = \lambda a, \lambda \neq 0$ which is closest to b. Then p is perpendicular to the vector $b - p$ which means $p^T(b-p) = 0$. We just need to find λ, so we solve $\lambda a^T(b - \lambda a) = 0$ and get $\lambda = \frac{a^Tb}{a^Ta}$.
Plugging in the numbers, we get $\lambda = \frac{10}{3}$, so the closest point is $\lambda a = \frac{10}{3}(1,1,1)$.
Equivalently the closest point to a is $\mu b = \frac{10}{36}b = \frac{10}{36}(2,4,4)$.

b) Prove that the trace of $P = aa^T/a^Ta$ always equals 1.

We just unfold $aa^T = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \begin{pmatrix} a_1 \dots a_n \end{pmatrix} = \begin{bmatrix} a_1^2 & a_1a_2 & \dots & a_1a_n \\ a_2a_1 & a_2^2 & \dots & a_2a_n \\ \vdots & \vdots & \ddots & \vdots \\ a_na_1 & a_na_2 & \dots & a_n^2 \end{bmatrix}$.
Also $a^Ta = \sum_i a_i^2$. Therefore the trace of P is $Tr(P) = Tr(aa^T/a^Ta) = \frac{a_1^2 + \dots a_n^2}{\sum_i a_i^2} = 1$.

c) Show that the length of Ax equals the length of A^Tx if $AA^T = A^TA$.

$||Ax||^2 = (Ax)^T(Ax) = x^TA^TAx = x^TAA^Tx = (A^Tx)^T(A^Tx) = ||A^Tx||^2$.

d) Which 2×2 matrix projects the x,y plane onto the line $x + y = 0$?

We are looking for the matrix $A \in \mathbb{R}^{2\times 2}$ that when multiplied with any vector $v = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$ gives us a vector u that is a *projection* of v on the line $x + y = 0$ or otherwise it is a vector $p = \lambda \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. This means that $Av = p$ and $p^T(v - p) = 0$.

Solving for $\lambda \neq 0$ we get

$$\begin{aligned} p^T(v-p) &= 0 \\ \lambda(1 \quad -1)\left(\begin{pmatrix} x \\ y \end{pmatrix} - \lambda\begin{pmatrix} 1 \\ -1 \end{pmatrix}\right) &= 0 \\ \lambda(x-y) - 2\lambda^2 &= 0 \\ \lambda &= \frac{1}{2}(x-y) \end{aligned}$$

$$\Rightarrow p = \frac{1}{2}(x-y)\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

So we have

$$\begin{aligned} Av &= p \\ \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} &= \frac{1}{2}(x-y)\begin{pmatrix} 1 \\ -1 \end{pmatrix} \\ &\Rightarrow \begin{cases} a_{11}x + a_{12}y & = \frac{1}{2}x - \frac{1}{2}y \\ a_{21}x + a_{22}y & = -\frac{1}{2}x + \frac{1}{2}y \end{cases} \end{aligned}$$

And since we have no other constraint for A, we use the obvious solution

$$A = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

Exercise 2: Determinants

a) If a square matrix A has determinant $\frac{1}{2}$, find $\det(2A)$, $\det(-A)$, $\det(A^2)$ and $\det(A^{-1})$.

$$\begin{aligned} \det(2A) = 2^n \det(A) = 2^n \frac{1}{2} &= 2^{n-1} \\ \det(-A) = (-1)^n \det(A) &= \pm\frac{1}{2} \\ \det(A^2) = \det(AA) = \det(A)\det(A) = (\frac{1}{2})^2 &= \frac{1}{4} \\ \det(A^{-1}) = \det(A)^{-1} = (\frac{1}{2})^{-1} &= 2 \end{aligned}$$

b) Find the determinants of

$$A = \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix}\begin{bmatrix} 2 & -1 & 2 \end{bmatrix} \quad , \quad U = \begin{bmatrix} 4 & 4 & 8 & 8 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 2 & 6 \\ 0 & 0 & 0 & 2 \end{bmatrix}, \; U^T \text{ and } U^{-1}$$

$$\det(A) = 0 \quad \text{(A has rank 1 so it is not invertible)}$$
$$\det(U) = \prod_{\lambda \in \{4,1,2,2\}} \lambda = 16 \quad \text{(product of the eigenvalues which lie on the diagonal on a triangular matri}$$
$$\det(U^T) = det(U) = 16$$
$$\det(U^{-1}) = \det(U)^{-1} = \frac{1}{16}$$

Exercise 3: Eigenvalues and Eigenvectors

a) Find the eigenvalues and eigenvectors of

$$A = \begin{bmatrix} 3 & 4 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 0 \end{bmatrix}$$, their traces and their determinants.

$$\det(A - \lambda I) = (3 - \lambda)(1 - \lambda)(-\lambda) = 0 \Rightarrow \lambda \in \{3, 1, 0\}$$

To find the eigenvectors we plug in the eigenvalues and solve the linear system $Ax = \lambda x$ for $x \neq 0$. The corresponding eigenvectors are then

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} \text{ and } \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix}$$

The trace and determinant are

$$\mathrm{Tr}(A) = 3 + 1 + 0 = 4$$
$$\det(A) = 0$$

For matrix B we have

$$\begin{aligned} \det(B - \lambda I) = (-\lambda)(2 - \lambda)(-\lambda) + 2(-2)(2 - \lambda) &= 0 \\ (\lambda^2 - 4)(2 - \lambda) &= 0 \\ (\lambda + 2)(\lambda - 2)(2 - \lambda) &= 0 \\ &\Rightarrow \lambda \in \{-2, 2, 2\} \end{aligned}$$

The corresponding eigenvectors are then

$$\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \text{ and } \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

The trace and determinant are

$$\text{Tr}(B) = 0 + 2 + 0 = 2$$
$$\det(B) = 2(0 - 4) = -8$$

Typically eigenvectors are normalized to have length 1 but any multiple of an eigenvector is also an eigenvector.

b) Using the characteristic polynomial, find the relationship between the trace, the determinants and the eigenvalues of any square matrix A.

We can factor the characteristic polynomial as a function of λ as

$$\det(A - \lambda I) = p(\lambda) = (-1)^n(\lambda - \lambda_1)\cdots(\lambda - \lambda_n) \tag{1}$$

where λ_i are the roots of the polynomial and the eigenvalues of A. We can simply set $\lambda = 0$ and find that

$$\det(A) = p(0) = (-1)^n(-\lambda_1)\cdots(-\lambda_n) = (-1)^n\prod_{i=1}^{n}(-\lambda_i) = (-1)^n\prod_{i=1}^{n}(-1)(\lambda_i)$$
$$= (-1)^n(-1)^n\prod_{i=1}^{n}\lambda_i = (-1)^{2n}\prod_{i=1}^{n}\lambda_i = \prod_{i=1}^{n}\lambda_i$$

So the determinant of a matrix is equal to the product of its eigenvalues.

Let us deal with the trace. Consider the 2×2 case

$$det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$
$$det(A - \lambda I) = \begin{vmatrix} a - \lambda & b \\ c & d - \lambda \end{vmatrix}$$
$$= (a - \lambda)(d - \lambda) - bc$$
$$= ad - bc - \lambda(a + d) + \lambda^2$$
$$= \lambda^2 - \lambda \cdot Tr(A) + \det(A)$$

Considering the $n \times n$ case and focusing on the diagonal, we find that

$$det(A - \lambda I) = (-\lambda)^n + (-\lambda)^{n-1} \cdot Tr(A) + \sum_{j=2}^{n-2}\beta_j\lambda^j + \det(A) \tag{2}$$

Comparing equations (1) and (2) we see that

$$Tr(A) = \lambda_1 + \lambda_2 + \ldots + \lambda_n = \sum_{i=1}^{n}\lambda_i \tag{3}$$

c) Diagonalize the unitary matrix V to reach $V = U\Lambda U^*$. All $|\lambda| = 1$. $V = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 & 1-i \\ 1+i & -1 \end{bmatrix}$

We have

$$\begin{aligned}
\det(V - \lambda I) &= (\frac{1}{\sqrt{3}} - \lambda)(-\frac{1}{\sqrt{3}} - \lambda) - \frac{1}{3}(1+i)(1-i) \\
&= (\frac{1}{\sqrt{3}} - \lambda)(-\frac{1}{\sqrt{3}} - \lambda) - \frac{2}{3} \\
&= -\frac{1}{3} + \lambda^2 - \frac{2}{3} \\
&= \lambda^2 - 1 = (\lambda - 1)(\lambda + 1)
\end{aligned}$$

Eigenvalues are $\lambda \in \{1, -1\}$ and corresponding eigenvectors are

$$x_1 = \frac{1}{\sqrt{1+2c^2}}\begin{pmatrix} 1 \\ c+ic \end{pmatrix} \quad \text{and} \quad x_2 = \frac{1}{\sqrt{1+2c^2}}\begin{pmatrix} -c+ic \\ 1 \end{pmatrix}$$

where $c = \frac{\sqrt{3}-1}{2}$.

Note that we could arrange the eigenvectors differently but since the matrix U is unitary, we have to keep the diagonal entries real. Now we can write matrix U as

$$U = \frac{1}{\sqrt{1+2c^2}}\begin{bmatrix} 1 & -c+ic \\ c+ic & 1 \end{bmatrix}$$

Therefore our decomposition can be written as

$$V = \frac{1}{1+2c^2}\begin{bmatrix} 1 & -c+ic \\ c+ic & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 1 & c-ic \\ -c-ic & 1 \end{bmatrix}$$

d) Suppose T is a 3×3 upper triangular matrix with entries t_{ij}. Compare the entries of T^*T and TT^*. Show that if they are equal, then T must be diagonal. (All normal triangular matrices are diagonal)

Let $T = \begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{bmatrix}$ with $a, b, c, d, e, f \in \mathbb{C}$.

Then

$$T^*T = \begin{bmatrix} \bar{a} & 0 & 0 \\ \bar{b} & \bar{d} & 0 \\ \bar{c} & \bar{e} & \bar{f} \end{bmatrix}\begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{bmatrix} = \begin{bmatrix} \bar{a}a & \bar{a}b & \bar{a}c \\ \bar{b}a & \bar{b}b + \bar{d}d & \bar{b}c + \bar{d}e \\ \bar{c}a & \bar{c}b + \bar{e}d & \bar{c}c + \bar{e}e + \bar{f}f \end{bmatrix}$$

$$TT^* = \begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{bmatrix}\begin{bmatrix} \bar{a} & 0 & 0 \\ \bar{b} & \bar{d} & 0 \\ \bar{c} & \bar{e} & \bar{f} \end{bmatrix} = \begin{bmatrix} a\bar{a} + b\bar{b} + c\bar{c} & b\bar{d} + c\bar{e} & c\bar{f} \\ d\bar{b} + e\bar{c} & d\bar{d} + e\bar{e} & e\bar{f} \\ f\bar{c} & f\bar{e} & f\bar{f} \end{bmatrix}$$

Now if $TT^* = T^*T$ we see from the diagonal entries that $-b\bar{b} = c\bar{c}$ and $\bar{b}b = e\bar{e}$. So, it must be that $b = c = e = 0$ and therefore T is diagonal.

Exercise 4: Singular Value Decomposition

a) Find the singular values and singular vectors of

$$A = \begin{bmatrix} 1 & 4 \\ 2 & 8 \end{bmatrix}$$

Eigenvalues of $A^T A$ are

$$\begin{aligned} \det(A^T A - \lambda I) &= \lambda(\lambda - 85) = 0 \\ &\Rightarrow \lambda \in \{0, 85\} \end{aligned}$$

Eigenvectors of $A^T A$ are $\begin{pmatrix} -4 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 4 \end{pmatrix}$ with norm $\sqrt{17}$.

Eigenvalues of AA^T are also $\lambda \in \{0, 85\}$

Eigenvectors of AA^T are $\begin{pmatrix} -2 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ with norm $\sqrt{5}$.

Therefore:

$$A = \frac{1}{\sqrt{17}} \begin{bmatrix} -4 & 1 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & \sqrt{85} \end{bmatrix} \frac{1}{\sqrt{5}} \begin{bmatrix} -2 & 1 \\ 1 & 2 \end{bmatrix}$$

b) Explain how UDV^T expresses A as a sum of r rank-1 matrices: $A = \sigma_1 u_1 v_1^T + \ldots + \sigma_r u_r v_r^T$

We see the factorization as

$$A = UDV^T = U(DV^T) = \begin{bmatrix} u_1 \ldots u_m \end{bmatrix} \left(\begin{bmatrix} \sigma_1 & & 0 & 0 & 0 & 0 \\ & \ddots & & & & 0 \\ 0 & & \sigma_r & 0 & 0 & 0 \\ 0 & \ldots & 0 & 0 & 0 & 0 \\ 0 & \ldots & 0 & 0 & \ddots & 0 \end{bmatrix} \begin{bmatrix} v_1^T \\ \vdots \\ v_n^T \end{bmatrix} \right)$$

$$= \begin{bmatrix} u_1 \ldots u_m \end{bmatrix} \begin{bmatrix} \sigma_1 v_1^T \\ \vdots \\ \sigma_r v_r^T \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \sigma_1 u_1 v_1^T + \ldots + \sigma_r u_r v_r^T + 0 \cdot u_{r+1} + \ldots + 0 \cdot u_m$$

Note that for the rank it holds $r \leq m$ and $r \leq n$.

c) If A changes to $4A$ what is the change in the SVD?

If $A = UDV^*$ then $4A = 4UDV^* = U(4D)V^*$. We apply the scaling to the singular values and leave the singular vectors normalized as they are.

What is the SVD for A^T and for A^{-1} ?

If $A = UDV^*$ then $A^T = (UDV^*)^T = VD^TU^T$ The singular values stay in the diagonal, but the dimensions of matrix D swap.

If $A = UDV^*$ then we can only compute the pseudoinverse $A^+ = (UDV^*)^+ = (V^*)^{-1}D^+U^{-1} = VD^+U^*$ Since U, V are unitary, their (conjugate) transpose is also their inverse. The reciprocals of the singular values are in the diagonal.

d) Find the SVD and the pseudoinverse of $A = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$ and $C = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$

The SVD of A will be $A = UDV^*$ where U is 1×1 meaning a scalar and since it is unitary it is 1, therefore $A = DV^*$.

$$A^TA = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \quad \text{and} \quad AA^T = 4$$

Then

$$\begin{aligned} \det(AA^T - \lambda I) &= 4 - \lambda = 0 \\ &\Rightarrow \lambda = 4 \end{aligned}$$

and

$$\begin{aligned} \det(A^TA - \lambda I) &= \ldots = \lambda^3(\lambda - 4) = 0 \\ &\Rightarrow \lambda \in \{0, 4\} \end{aligned}$$

For $\lambda = 4$ we get one eigenvector $v_1 = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$. For $\lambda = 0$ we get three eigenvectors with only one constraint, that the sum of their entries is zero. We choose them to be orthogonal to each other and normalize them, so that matrix V is indeed unitary.

$$v_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} v_3 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} v_4 = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}$$

AA^T has one eigenvalue $\lambda = 4$, therefore $\sigma = 2$ and $D = \begin{bmatrix} 2 & 0 & 0 & 0 \end{bmatrix}$ since A has rank 1.

We now can write the SVD of A:

$$A = UDV^* = \begin{bmatrix}1\end{bmatrix} \begin{bmatrix}2 & 0 & 0 & 0\end{bmatrix} c \begin{bmatrix} c & c & c & c \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ c & c & -c & -c \end{bmatrix}$$

where $c = \frac{1}{\sqrt{2}}$.

The pseudoinverse of A is then

$$A^+ = VDU^* = c \begin{bmatrix} c & 1 & 0 & c \\ c & -1 & 0 & c \\ c & 0 & 1 & -c \\ c & 0 & -1 & -c \end{bmatrix} \begin{bmatrix} \frac{1}{2} \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix}1\end{bmatrix} = \frac{c^2}{2} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

For B we have

$$B = UDV^* = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and therefore pseudoinverse

$$B^+ = VD^+U^* = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 0 \end{bmatrix}$$

Finally, for C we have

$$C = UDV^* = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{2} & 0 \\ 0 & 0 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

and therefore pseudoinverse

$$C^+ = VD^+U^* = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & 0 \\ \frac{1}{2} & 0 \end{bmatrix}$$