233 Computational Techniques

Problem Sheet for Tutorial 4

Problem 1
In the standard basis of $\mathbb{R}^2$, let the linear map $f : \mathbb{R}^2 \to \mathbb{R}^2$ have matrix representation

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

Find the eigenvalues and eigenvectors of $\boldsymbol{A}$. Hence find the basis with respect to which $\boldsymbol{A}$ is a diagonal matrix and find the matrix for this change of basis.

Problem 2
Find the singular value decomposition of the matrix.

$$\boldsymbol{A} = \begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix}.$$

Problem 3
Show that, for any matrix $\boldsymbol{A} \in \mathbb{R}^{m\times m}$, if $\boldsymbol{v}$ is an eigenvector of $\boldsymbol{A}^T\boldsymbol{A}$ with eigenvalue $\lambda \neq 0$, then $\boldsymbol{A}\boldsymbol{v}$ is an eigenvector of $\boldsymbol{A}\boldsymbol{A}^T$ with the same eigenvalue. (Why do we need $\lambda \neq 0$ here?) Show that if $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are orthogonal eigenvectors of $\boldsymbol{A}^T\boldsymbol{A}$, then $\boldsymbol{A}\boldsymbol{v}_1$ and $\boldsymbol{A}\boldsymbol{v}_2$ are orthogonal. State and prove a similar result for eigenvectors of $\boldsymbol{A}\boldsymbol{A}^T$. Deduce that for any matrix $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, the two matrices $\boldsymbol{A}^T\boldsymbol{A}$ and $\boldsymbol{A}\boldsymbol{A}^T$ have the same set of non-zero eigenvalues.

Problem 4

(i) Show that an orthogonal transformation preserves the angle between any two vectors.

(ii) Show that an orthogonal transformation preserves the ℓ_2 norm of a vector. Hence, use the SVD representation of any matrix $\boldsymbol{A}$ to show that the $\|\boldsymbol{A}\|_2 := \sup_{\|x\|_2=1} \|\boldsymbol{A}\boldsymbol{x}\|_2$ is equal to σ_1 the largest singular value of $\boldsymbol{A}$.

Problem 5
The purpose of this exercise is to show you an application of eigenvalues and eigenvectors to a topic which, at first glance, might seem totally unrelated: the *Fibonacci series*.
Recall (from the 1st year PPT classes) that the series is defined by $x_0 := 0$, $x_1 := 1$ and

$$x_{n+1} := x_n + x_{n-1} \tag{1}$$

for $n \geq 1$. This formula is *recursive*, that is, in order to find x_n for higher values of n, you have to know (or compute) the values for smaller n.
In many situations recursive formulae are not good enough, for instance if one wants to know how x_n grows with n. In this exercise you can find a formula for x_n which is *non-recursive* in the sense that it gives x_n as a *function of the index* n rather than as a function of previously computed values. Eigenvalues and -vectors are a good tool for this. Here is how to do it:
(a) Express (1) as a vector equation of the form

$$\begin{bmatrix} x_{n+1} \\ x_n \end{bmatrix} = \boldsymbol{A} \begin{bmatrix} x_n \\ x_{n-1} \end{bmatrix} \tag{2}$$

for some 2×2 matrix $\boldsymbol{A}$. This transforms the original series into a series of two-dimensional vectors.
(b) By recursive application of (2), express $[x_{n+1}, x_n]^T$ as a power of $\boldsymbol{A}$ times the "initial" vector (which one)?
(c) Now, find eigenvalues λ_i and eigenvectors $\boldsymbol{u}_i$ of $\boldsymbol{A}$. (Here the $\boldsymbol{u}_i$ need not be normalized.)
(d) Express the initial vector as a linear combination of the eigenvectors of $\boldsymbol{A}$.
(e) Use the results of *(b)–(d)* and the relation $\boldsymbol{A}\boldsymbol{u}_i = \lambda_i \boldsymbol{u}_i$ to find the vector $[x_{n+1}, x_n]$—and hence x_n itself—as a function of n alone.
(f) Test your formula for $n = 0, \ldots, 4$.

Problem 6
Using the fact that linear independence of the columns (or rows) of a matrix $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ is invariant under elementary row or column operations, as proved in the notes, show that the column rank and the row rank of a matrix is invariant under elementary row or column operations.
Hint: Consider the column rank of $\boldsymbol{A}$. (i) Elementary column operations: For the elementary operation of swapping two columns or multiplying one by a non-zero real number the assertion is clear. Consider the elementary operation of subtracting $\lambda\boldsymbol{a}_2$ from $\boldsymbol{a}_1$. Take a set S of maximally independent column vectors of the matrix and consider the four cases where $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ belong or do not belong to this set. (ii) Elementary row operations: Consider any elementary row operation on the set S of a maximally independent column vectors of the matrix.

Solutions

Problem 1

The eigenvalues are 0 with eigenvector $\boldsymbol{v}_1 = \begin{bmatrix}-1\\2\end{bmatrix}$ and 5 with eigenvector $\boldsymbol{v}_2 = \begin{bmatrix}2\\1\end{bmatrix}$. If we take $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ as the new basis, then the matrix representing f would be $\begin{bmatrix}0&0\\0&5\end{bmatrix}$. The matrix for change of basis is $\mathbf{B} = [\boldsymbol{v}_1, \boldsymbol{v}_2]$.

Problem 2

We find the SVD of the matrix as $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^T$. First we find the eigenvalues and

eigenvectors of

$$\boldsymbol{A}^T\boldsymbol{A} = \begin{bmatrix}4&-3\\4&3\end{bmatrix}\begin{bmatrix}4&4\\-3&3\end{bmatrix} = \begin{bmatrix}25&7\\7&25\end{bmatrix}.$$

The eigenvalues are 32 with normalised eigenvector $\boldsymbol{v}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}$ and 18 with normalised eigenvector $\boldsymbol{v}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix}$. Thus, we have $\sigma_1 = \sqrt{3}2 = 4\sqrt{2}$ and $\sigma_2 = \sqrt{1}8 = 3\sqrt{2}$ with

$$\boldsymbol{V} = \left[\frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix}\right] = \frac{1}{\sqrt{2}}\begin{bmatrix}1&1\\1&-1\end{bmatrix}$$

Next we put

$$\boldsymbol{u}_1 = \frac{1}{\sigma_1}\boldsymbol{A}\boldsymbol{v}_1 = \begin{bmatrix}4&4\\-3&3\end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix} = \frac{1}{4\sqrt{2}}\frac{1}{\sqrt{2}}\begin{bmatrix}8\\0\end{bmatrix} = \begin{bmatrix}1\\0\end{bmatrix}$$

$$\boldsymbol{u}_2 = \frac{1}{\sigma_2}\boldsymbol{A}\boldsymbol{v}_2 = \frac{1}{3\sqrt{2}}\begin{bmatrix}4&4\\-3&3\end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix} = \frac{1}{3\sqrt{2}}\frac{1}{\sqrt{2}}\begin{bmatrix}0\\-6\end{bmatrix} = \begin{bmatrix}0\\-1\end{bmatrix}$$

Putting $\boldsymbol{U} = [\boldsymbol{u}_1, \boldsymbol{u}_2]$ and $\boldsymbol{S} = \text{diag}(4\sqrt{2}, 3\sqrt{2})$, a simple calculation shows that we have

$$\boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^T = \boldsymbol{A}.$$

Problem 3

From $\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{v}$, it follows that $\boldsymbol{A}\boldsymbol{v}$ is not the zero vector since otherwise, as $\lambda \neq 0$, we obtain $\boldsymbol{v} = \boldsymbol{0}$, which is a contradiction becasue $\boldsymbol{v}$ is an eigenvector. By pre-multiplying both sides of $\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{v}$ with $\boldsymbol{A}$, we get: $\boldsymbol{A}\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{A}\boldsymbol{v}$ or $\boldsymbol{A}\boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{v}) = \lambda\boldsymbol{A}\boldsymbol{v}$ as desired. Also if $\boldsymbol{v}_2$ and $\boldsymbol{v}_2$ are orthogonal eigenvectors of $\boldsymbol{A}^T\boldsymbol{A}$ with eigenvalues λ_1 and λ_2 respectively, then $(\boldsymbol{A}\boldsymbol{v}_2)^T\boldsymbol{A}\boldsymbol{v}_1 = \boldsymbol{v}_2^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{v}_1 = \lambda_1\boldsymbol{v}_2^T\boldsymbol{v}_1 = 0$. We obtain the same results if we start with an eigenvector $\boldsymbol{u}$ of $\boldsymbol{A}\boldsymbol{A}^T$ and consider $\boldsymbol{A}^T\boldsymbol{u}$, which will be an eigenvector of $\boldsymbol{A}^T\boldsymbol{A}$.

Problem 4

(i) If $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^m$ are unit vectors and $\boldsymbol{U} \in \mathbb{R}^{m\times m}$ is an orthogonal matrix then the cosine of the angle between them is $\boldsymbol{x}\cdot\boldsymbol{y}$. But $(\boldsymbol{U}\boldsymbol{x})\cdot(\boldsymbol{U}\boldsymbol{y}) = (\boldsymbol{U}\boldsymbol{x})^T(\boldsymbol{U}\boldsymbol{y}) = (\boldsymbol{x}^T\boldsymbol{U}^T)(\boldsymbol{U}\boldsymbol{y}) = \boldsymbol{x}^T(\boldsymbol{U}^T\boldsymbol{U})\boldsymbol{y} = \boldsymbol{x}^T\boldsymbol{y} = \boldsymbol{x}\cdot\boldsymbol{y}$, since $\boldsymbol{U}^T\boldsymbol{U} = \boldsymbol{I}$.
(ii)If $\boldsymbol{x} \in \mathbb{R}^m$ and $\boldsymbol{U} \in \mathbb{R}^{m\times m}$ is an orthogonal matrix then we have: $(\|\boldsymbol{U}\boldsymbol{x}\|_2)^2 = (\boldsymbol{U}\boldsymbol{x})^T\boldsymbol{U}\boldsymbol{x} = (\boldsymbol{x}^T\boldsymbol{U}^T)\boldsymbol{U}\boldsymbol{x} = \boldsymbol{x}^T(\boldsymbol{U}^T\boldsymbol{U})\boldsymbol{x} = \boldsymbol{x}^T\boldsymbol{x} = (\|\boldsymbol{x}\|_2)^2$, since $\boldsymbol{U}^T\boldsymbol{U} = \boldsymbol{I}$.
Now let $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^T$ be the SVD of $\boldsymbol{A}$. The vector $\boldsymbol{x} \in \mathbb{R}^m$ is in one-to-one correspondence with $\boldsymbol{y} = \boldsymbol{V}^T\boldsymbol{x} \in \mathbb{R}^m$ and have the same ℓ_2 norm. Then

$$\|\boldsymbol{A}\|_2 := \sup_{\|x\|_2=1} \|\boldsymbol{A}\boldsymbol{x}\|_2 = \sup_{\|x\|_2=1} \|\boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^T\boldsymbol{x}\|_2$$

$$= \sup_{\|y\|_2=1} \|\boldsymbol{U}\boldsymbol{S}\boldsymbol{y}\|_2 = \sup_{\|y\|_2=1} \|\boldsymbol{S}\boldsymbol{y}\|_2,$$

since the orthogonal matrix $\boldsymbol{U}$ preserves the ℓ_2 norm.
Thus,

$$(\|\boldsymbol{A}\|_2)^2 = \sup_{\|y\|_2=1} \{(\sigma_1 y_1)^2 + (\sigma_2 y_2)^2 + \cdots + (\sigma_p y_p)^2\},$$

where $p = \min(m, n)$. But $\sigma_1^2 y_1^2 + \sigma_2^2 y_2^2 + \cdots + \sigma_p^2 y_p^2 \leq \sigma_1^2(y_1^2 + y_2^2 + \cdots y_p^2) \leq \sigma_1^2$ for all $\boldsymbol{y} \in \mathbb{R}^m$ with $\|\boldsymbol{y}\|_2 = 1$ and for $\boldsymbol{y}^T = (1, 0, 0, \cdots, 0)$ we have

$$(\sigma_1 y_1)^2 + (\sigma_2)^2 y_2^2 + \cdots + (\sigma_p)^2 y_p^2 = \sigma_1^2$$

which implies $\|\boldsymbol{A}\|_2 = \sigma_1$.

Problem 5

(a) The matrix in (2) is

$$\boldsymbol{A} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}.$$

(b) The initial vector is $[x_1, x_0]^T = [1, 0]^T$, and

$$\begin{bmatrix} x_{n+1} \\ x_n \end{bmatrix} = \boldsymbol{A}^n \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

(c) The characteristic polynomial of $\boldsymbol{A}$ is

$$\det(\boldsymbol{A} - \lambda \boldsymbol{I}_2) = \det \begin{bmatrix} 1-\lambda & 1 \\ 1 & -\lambda \end{bmatrix} = -\lambda(1-\lambda) - 1 = \lambda^2 - \lambda - 1 ,$$

with zeros $\lambda_1 = (1+\sqrt{5})/2$ and $\lambda_2 = (1-\sqrt{5})/2$; these are the eigenvalues of $\boldsymbol{A}$. A corresponding choice of eigenvectors is

$$\boldsymbol{u}_1 = \begin{bmatrix} \lambda_1 \\ 1 \end{bmatrix} , \qquad \boldsymbol{u}_2 = \begin{bmatrix} \lambda_2 \\ 1 \end{bmatrix} .$$

They are not normalized – this is not necessary here as we do not need the explicit orthogonal matrix from the spectral decomposition.
(d)

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{\boldsymbol{u}_1 - \boldsymbol{u}_2}{\sqrt{5}} .$$

(e) Multiplying both sides of the last equation by $\boldsymbol{A}^n$ gives

$$\begin{bmatrix} x_{n+1} \\ x_n \end{bmatrix} = \frac{\boldsymbol{A}^n \boldsymbol{u}_1 - \boldsymbol{A}^n \boldsymbol{u}_2}{\sqrt{5}} = \frac{1}{\sqrt{5}} \left\{ \lambda_1^n \begin{bmatrix} \lambda_1 \\ 1 \end{bmatrix} - \lambda_2^n \begin{bmatrix} \lambda_2 \\ 1 \end{bmatrix} \right\} .$$

Here the second component gives

$$x_n = \frac{\lambda_1^n - \lambda_2^n}{\sqrt{5}} = \frac{\{(1+\sqrt{5})/2\}^n - \{(1-\sqrt{5})/2\}^n}{\sqrt{5}} . \qquad (3)$$

(f) Thus:

$$\begin{aligned} x_0 &= 0 , \\ x_1 &= \frac{(1+\sqrt{5})/2 - (1-\sqrt{5})/2}{\sqrt{5}} = 1 , \\ x_2 &= \frac{(1+2\sqrt{5}+5)/4 - (1-2\sqrt{5}+5)/4}{\sqrt{5}} = 1 , \\ x_3 &= \frac{(1+3\sqrt{5}+15+5\sqrt{5})/8 - (1-3\sqrt{5}+15-5\sqrt{5})/8}{\sqrt{5}} = 2 , \\ x_4 &= \frac{(1+4\sqrt{5}+30+20\sqrt{5}+25)/16 - (1-4\sqrt{5}+30-20\sqrt{5}+25)/16}{\sqrt{5}} = 3 \end{aligned}$$

in agreement with (1).
Obviously the recursive formula is better for small values of n as it avoids the "detour" into the real numbers. However for large n, (3) with real arithmetic can be much faster than (1) or (2) with integer arithmetic.

Problem 6

By symmetry we only to prove the assertion for the column rank of $\boldsymbol{A}$ since then the assertion about the row rank follows by considering $\boldsymbol{A}^T$. We use the following fact proved in the notes:

$$(*) \qquad \text{Linear independence of the columns (or rows) of a matrix } \boldsymbol{A} \in \mathbb{R}^{m\times n}$$
$$\text{is invariant under elementary row or column operations.}$$

We consider the two cases of elementary column operation and elementary row operations in turn.
(a) Elementary column operations: on $\boldsymbol{A}$: For the elementary operation of swapping two columns or multiplying one by a non-zero real number the assertion is clear. Consider the elementary operation of subtracting $\lambda\boldsymbol{a}_2$ from $\boldsymbol{a}_1$, where we assume that $\boldsymbol{a}_2 \neq \boldsymbol{0}$. Take a set S of maximally independent column vectors of the matrix, with say r vectors, which correspond to a submatrix of $\boldsymbol{A}$. Let $\boldsymbol{A}'$ be the result of the elementary operation on $\boldsymbol{A}$ and S' be the result of the elementary operation on S. Consider the four cases where $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ belong or do not belong to this set:
(i) $\boldsymbol{a}_1 \in S$ and $\boldsymbol{a}_2 \in S$: By (*) the elementary operation preserves the linear independence of the columns in S. Thus, the vectors in S' are linearly independent and since all column vectors not in S can be expressed in terms of S and therefore in terms of S', it follows that S' is a maximally linearly independent set of column vectors of $\boldsymbol{A}'$. Thus, the column rank is preserved.
(ii) $\boldsymbol{a}_1 \notin S$ and $\boldsymbol{a}_2 \notin S$: In this case, we have $S = S'$; furthermore $\boldsymbol{a}_1 - \lambda\boldsymbol{a}_2$ can be expressed in terms of column vectors in S as S is a set of maximally independent column vectors of $\boldsymbol{A}$. Thus, $S' = S$ is still a set of maximally independent column vectors of $\boldsymbol{A}'$.
(iii) $\boldsymbol{a}_1 \notin S$ and $\boldsymbol{a}_2 \in S$: Here we have $S' = S$. Since $\boldsymbol{a}_1 - \lambda\boldsymbol{a}_2$ can be expressed as a linear combination of vectors in S, the set $S' = S$ remains a set of maximally linearly independent column vectors again.
(iv) $\boldsymbol{a}_1 \in S$ and $\boldsymbol{a}_2 \notin S$: There are two sub-cases here.

Sub-case (1): Suppose S' has rank r. Let $s \subset \{1, 2, \ldots, n\}$ be such that $\boldsymbol{a}_i \in S$ iff $i \in s$ and let $\boldsymbol{a}_0 = \boldsymbol{a}_1 - \lambda\boldsymbol{a}_2$ and $s' = \{0\} \cup (s \setminus \{1\})$. We assert that S' is a set of maximally linearly independent vectors. To show this, it is sufficient to show that we can express $\boldsymbol{a}_1$ as a linear combination of vectors in S'. Since $\boldsymbol{a}_1 = (\boldsymbol{a}_1 - \lambda\boldsymbol{a}_2) + \lambda\boldsymbol{a}_2$ and $\boldsymbol{a}_1 - \lambda\boldsymbol{a}_2 \in S'$, this latter claim follows if we show that $\boldsymbol{a}_2$ can be expressed as a linear combination of elements in S'. To obtain this linear combination, we need to find $y_i \in \mathbb{R}$ for $i \in s'$ such that $\boldsymbol{a}_2 = \sum_{i\in s'} y_i\boldsymbol{a}_i = y_0(\boldsymbol{a}_1 - \lambda\boldsymbol{a}_2) + \sum_{i\in s'\setminus\{0\}} \boldsymbol{a}_i$.
Since S is maximally linearly independent, there exist $x_i \in \mathbb{R}$ for $i \in s$, with

$$\boldsymbol{a}_2 = \sum_{i\in s} x_i\boldsymbol{a}_i \tag{4}$$

which can be rewritten as

$$(1 - x_1\lambda)\boldsymbol{a}_2 = x_1(\boldsymbol{a}_1 - \lambda\boldsymbol{a}_2) + \sum_{i \in s\setminus\{1\}} x_i\boldsymbol{a}_i. \tag{5}$$

Note that $1 - x_1\lambda \neq 0$ since otherwise, the LHS of Equation (5) vanishes and by linear independence of vectors in S', from the RHS of Equation (5) we get $x_i = 0$ for all $i \in s$; but this implies, by Equation (4) that $\boldsymbol{a}_2 = 0$ contrary to our assumption.
Thus, we obtain

$$\boldsymbol{a}_2 = \frac{1}{1 - x_1\lambda}(x_1(\boldsymbol{a}_1 - \lambda\boldsymbol{a}_2) + \sum_{i \in s\setminus\{1\}} x_i\boldsymbol{a}_i) = \frac{1}{1 - x_1\lambda}(x_1(\boldsymbol{a}_1 - \lambda\boldsymbol{a}_2) + \sum_{i \in s'\setminus\{0\}} x_i\boldsymbol{a}_i),$$

as required. This completes case (1).

Sub-case (2): Suppose S' has rank $r-1$. This means that $\boldsymbol{a}_1 - \lambda\boldsymbol{a}_2$ can be expressed as a linear combination of vectors in $S \setminus \boldsymbol{a}_1$. Thus, $\boldsymbol{a}_1 = (\boldsymbol{a}_1 - \lambda\boldsymbol{a}_2) + \lambda\boldsymbol{a}_2$ can be expressed as a linear combination of vectors in $T := (S \setminus \boldsymbol{a}_1) \cup \boldsymbol{a}_2$. Thus T is a set of maximally linearly independent column vectors with r vectors and since $T' = T$ it follows that the column rank is preserved in this case as well.

Therefore the maximal number of linearly independent column vectors, i.e., the column rank, is preserved under any elementary column operation.

(b) Elementary row operations: Consider any elementary row operation on the set S of maximally independent column vectors of $\boldsymbol{A}$ as a sub-matrix of $\boldsymbol{A}$ with say r vectors. Then, by (*), under the elementary row operation the vectors in S will remain linearly independent. Moreover any set of $r+1$ column vectors of $\boldsymbol{A}$ will be linearly dependent and thus, by (*) again, will remain linearly dependent vectors. Therefore the maximal number of linearly independent column vectors, i.e., the column rank, is preserved under any elementary row operation.