Elements of Mathematics

Exercise Sheet 6

Submission due date: **07.12.2021, 10:15h**

THEORY

1 Thin and Fat Full Rank Matrices

Answer the following questions without using the SVD. Instead, exploit the orthogonality relation between the four fundamental subspaces and the dimension formula.

1. What is the orthogonal complement of $\{0\}$ and $\mathbb{R}^n$ in $\mathbb{R}^n$, respectively?
2. Give an example for a matrix $C \in \mathbb{R}^{m\times n}$ with nullity$(C) = 0$ (injective).
   (a) What do we know about the order relation between m and n? (which one is greater or equal than the other?)
   (b) What do we know about the columns of C?
   (c) Let $b \in \mathrm{Im}(C)$. Can we find two distinct $x_1 \neq x_2 \in \mathbb{R}^n$ such that $Cx_1 = b = Cx_2$? Explain your answer.
   (d) What do we know about the matrix $C^\top C \in \mathbb{R}^{n\times n}$?
3. Give an example for a matrix $A \in \mathbb{R}^{m\times n}$ with $n > m$ and rank$(A) < m$.
4. Give an example for a matrix $R \in \mathbb{R}^{m\times n}$ with rank$(R) = m$ (surjective).
   (a) What do we know about the order relation between m and n?
   (b) What do we know about the rows of R?
   (c) What do we know about the matrix $RR^\top \in \mathbb{R}^{m\times m}$?
   (d) Let $b \in \mathbb{R}^m$. Can we find an $x \in \mathbb{R}^n$ such that $Rx = b$? Explain your answer and give an example for your matrix.
5. Give an an example for a matrix $A \in \mathbb{R}^{n\times n}$ with rank$(A) = n$ (invertible).
   (a) What do we know about the dimensions of $\ker(A)$, $\ker(A^\top)$ and $\mathrm{Im}(A^\top)$?
   (b) What do we know about the columns and rows of A?
   (c) Let $b \in \mathbb{R}^n$. Can we find an *unique* $x \in \mathbb{R}^n$ such that $Ax = b$? Explain your answer and give an example for your matrix.
6. **Bonus*:** Give an an example for a matrix $A \in \mathbb{R}^{m\times n}$ with rank$(A) =$ rank$(A^\top)$ and nullity$(A) \neq$ nullity$(A^\top)$.

(18 Points)

Solution:

Let us recall the dimension formula here

$$n = \text{rank}(A) + \text{nullity}(A),$$
$$m = \text{rank}(A^\top) + \text{nullity}(A^\top).$$

1. We have
$$\{0\}^\perp = \{x \in \mathbb{R}^n : x^\top v = 0 \ \ \forall v \in \{0\}\} = \{x \in \mathbb{R}^n : x^\top 0 = 0\} = \mathbb{R}^n$$
and
$$(\mathbb{R}^n)^\perp = \{x \in \mathbb{R}^n : x^\top y = 0 \ \ \forall y \in \mathbb{R}^n\} = \{x \in \mathbb{R}^n : \ker(x^\top) = \{0\}\} = \{0\}.$$

   [Not part of the exercise:]
   Together with $(U^\perp)^\perp = U$ this gives in general
$$\begin{aligned} \mathbb{R}^m = \mathrm{Im}(A) = \ker(A^\top)^\perp &\iff \ker(A^\top) = \{0\}, \quad \text{or} \\ \text{rank}(A) = m &\iff \text{nullity}(A^\top) = 0, \quad \text{or} \\ A \ \text{ surjective} &\iff A^\top \ \text{ injective.} \end{aligned}$$

   Analogously for the transpose
$$\begin{aligned} \mathbb{R}^n = \mathrm{Im}(A^\top) = \ker(A)^\perp &\iff \ker(A) = \{0\}, \quad \text{or} \\ \text{rank}(A^\top) = n &\iff \text{nullity}(A) = 0, \quad \text{or} \\ A^\top \ \text{ surjective} &\iff A \ \text{ injective.} \end{aligned}$$

2. Example:
$$C = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

   (a) By dimension formula we know
$$m \geq \dim(\mathrm{Im}(C)) = \text{rank}(C) = n - \text{nullity}(C) = n.$$

   Only square or thin matrices can be injective!

   (b) Since $\text{nullity}(C) = 0$ we know $\ker(C) = \{0\}$ and thus the n columns of C are independent (in fact, only the zero combination gives the zero vector).

   (c) No (because f_C injective). Recall the proof: Assume yes, then $Cx_1 = b = Cx_2 \iff C(x_1 - x_2) = 0$, where $x_1 - x_2 \neq 0$ due to $x_1 \neq x_2$. This contradicts the fact that $\ker(C) = \{0\}$.

   Independent columns assure that a solution to $Cx = b$ is unique (if it exists).

   (d) Since $\ker(C^\top C) = \ker(C) = \{0\}$, we have that the $(n \times n)$-matrix $C^\top C$ is invertible.

3. Example: Take for two nonzero vectors $u \in \mathbb{R}^m, v \in \mathbb{R}^n$ with $n > m > 1$ the outer product
$$A := uv^\top,$$
so that each column is a scaling of u and thus $\text{rank}(A) = 1 < m$.

   Simply many columns are not enough for surjectivity!
   We need independence to get a "larger" subspace.

4. Example: $R := C^\top$.

(a) From above recall: $\text{rank}(R) = m \iff \text{nullity}(R^\top) = 0$. Now by dimension formula we get

$$n \geq \dim(\text{Im}(R^\top)) = \text{rank}(R^\top) = m - \text{nullity}(R^\top) = m.$$

Only square or fat matrices can be surjective!

(b) Again, by $\text{nullity}(R^\top) = 0$, they are independent.

(c) Since $\ker(RR^\top) = \ker(R^\top) = \{0\}$, it's invertible.

(d) Yes, since $\text{Im}(R) = \mathbb{R}^m$, any $b \in \mathbb{R}^m$ is of the form $Rx = b$ for some $x \in \mathbb{R}^n$.

Independent rows assure that a solution to $Rx = b$ exists.

5. (a) By dimension formula

$$\text{nullity}(A) = n - \text{rank}(A) = 0,$$

Then using from above $\text{rank}(A^\top) = n \iff \text{nullity}(A) = 0$, we find $\text{rank}(A^\top) = n$, and therefore also

$$\text{nullity}(A^\top) = n - \text{rank}(A^\top) = 0.$$

(b) Since $\ker(A) = \{0\} = \ker(A^\top)$ the n rows and the n columns are independent.

(c) Yes, because from above we know: Independent rows give existence and independent columns uniqueness.

6. Take $A = C$ with C from above. Then

$$\text{rank}(A) = 2 = \text{rank}(A^\top).$$

However

$$\text{nullity}(C) = 0 \neq 1 = \text{nullity}(C^\top).$$

We will learn below:
Always $\text{rank}(A) = \text{rank}(A^\top)$,
but a similar result is not true in general for $\text{nullity}(A)$.

2 Equivalent Definitions of the Matrix Rank

Let $A \in \mathbb{R}^{m \times n}$, then show that the following statements are equivalent:

i) The maximum number of linearly independent columns of A is r.

ii) The dimension of the image of A is r, i.e., $\dim(\text{Im}(A)) = r$.

iii) The number of positive singular values of A is r.

Hint: You can use the SVD $A = U\Sigma V^\top$ and Lemma 2.29.

As a consequence (since $A^\top = V\Sigma^\top U^\top$), we find

$$\text{rank}(A) = r = \text{rank}(A^\top),$$

so that the dimension formulas for A and $A^\top$ read as

$$n = \text{rank}(A) + \text{nullity}(A) = \text{rank}(A^\top) + \text{nullity}(A),$$
$$m = \text{rank}(A^\top) + \text{nullity}(A^\top) = \text{rank}(A) + \text{nullity}(A^\top).$$

(6 Points)

Solution:

i) $\Rightarrow$ ii): Pick r independent columns of $A = [a_1, \ldots, a_n]$, say $a_{i_1}, \ldots, a_{i_r}$. Then by i) any other column of A can be written as linear combination of $a_{i_1}, \ldots, a_{i_r}$, so that

$$\mathrm{Im}(A) = \mathrm{span}(a_1, \ldots, a_n) = \mathrm{span}(a_{i_1}, \ldots, a_{i_r}).$$

Thus the $a_{i_1}, \ldots, a_{i_r}$ are a basis of length r for $\mathrm{Im}(A)$, implying ii) by the definition of "dimension".

ii) $\Rightarrow$ i): Let $\dim(\mathrm{Im}(A)) = r$. Since any basis of a subspace has the same length, any basis of $\mathrm{Im}(A)$ has length r.
Now assume A has more than r independent columns. Then by the reasoning from above, these columns would yield another basis of $\mathrm{Im}(A)$ but with length $> r$, which would contradict the fact that any basis has length r. Therefore implying i): the maximum number of independent columns in A is r.

ii) $\Leftrightarrow$ iii) We show that

$$\dim(\mathrm{Im}(A)) = \text{number of positive singular values of } A.$$

Let us consider the reduced SVD $A = U_r \Sigma_r V_r^\top$ where r denotes the number of (positive) singular values, Σ_r is an invertible diagonal matrix and U_r and $V_r^\top$ have independent (even orthonormal) columns and rows, respectively. Thus

$$\mathrm{Im}(A) = \mathrm{Im}(U_r \Sigma_r V_r^\top).$$

In order to put Lemma 2.29 into position we show that $\Sigma_r V_r^\top$ is surjective, i.e., $\mathrm{rank}(\Sigma_r V_r^\top) = r$. This easily follows from the fact that $(\Sigma_r V_r^\top)^\top = V_r \Sigma_r$ has independent columns and thus

$$\mathrm{nullity}(V_r \Sigma_r) = 0 \Leftrightarrow \mathrm{rank}(\Sigma_r V_r^\top) = r.$$

Therefore by the mentioned Lemma we obtain

$$\mathrm{Im}(A) = \mathrm{Im}(U_r \Sigma_r V_r^\top) = \mathrm{Im}(U_r).$$

Since the columns of U_r are independent they are a basis of length r for $\mathrm{Im}(A)$, which implies $\dim(\mathrm{Im}(A)) = r$=number of (positive) singular values.

PROGRAMMING

---

3 The QR Algorithm

The QR-**Algorithmn** is an eigenvalue algorithm. Thus, it is used to compute eigenvalues and eigenvectors of a matrix $A \in \mathbb{R}^{n \times n}$. It produces a sequence of matrices $(A_k)_{k \in \mathbb{N}}$. All A_k are similar to A and thus have the same eigenvalues. The iteration is defined as follows:

$$\begin{aligned}
&A_0 = A \in \mathbb{R}^{n \times n} \\
&\textbf{for } k = 0, \ldots, \infty: \\
&\qquad Q_{k+1} R_{k+1} := A_k \quad (QR \text{ decomposition}) \\
&\qquad A_{k+1} := R_{k+1} Q_{k+1}
\end{aligned}$$

If the absolute values of the eigenvalues of A are distinct, one can show that $A_\infty := \lim_{k \to \infty} A_k$ is a diagonal matrix. In this case, the eigenvalues of A are the diagonal elements of A_∞.

Task:

1. Implement the QR eigenvalue algorithm as a function `eig(A,m)`. The function shall take as input a matrix $A \in \mathbb{R}^{n \times n}$ and a maximum iteration number $m \in \mathbb{N}$. It shall return the diagonal of the last iterate A_m. For the QR decomposition you can use the Gram-Schmidt algorithm from previous sheets which we have implemented as a function `QR(A)` or an appropriate Python routine.

   Hint: You can access the diagonal of a `numpy.array` by `A.diagonal()`.

2. Test your algorithm on a random matrix $A \in \mathbb{R}^{n\times n}$. In order to generate such a random matrix use the following code snippet:

```
def A_gen(n):
   from numpy as np
   from scipy.linalg import qr
   A = np.random.rand(n,n)
   Q, R = qr(A)
   Lambda = np.diag(np.arange(1,n+1))
   A = Q @ (Lambda @ Q.T)
return A
```

3. Find a routine in Scipy to compute the eigenvalues and -vectors of a matrix. Test the routine on multiple examples, especially for higher dimensions. Compare to your algorithm.

(6 Points)

Solution:

```python
import numpy as np

def qr_factor(A):
    """
    Computes a QR-decomposition of a (mxn)-matrix with m>=n via Gram-Schmidt.

    Parameters
    ----------
    A : (mxn) matrix with m>=n

    Returns
    -------
    Q : (mxn) with orthonormal columns
    R : (nxn) upper triangular matrix
    """
    m, n = A.shape
    R = np.zeros((n,n))
    Q = np.zeros((m,n))

    R[0,0] = np.linalg.norm(A[:,0])
    Q[:,0] = A[:,0]/R[0,0]

    for k in range(1,n):
        for l in range(0, k):
            R[l,k] = A[:,k] @ Q[:,l]
        q = A[:,k] - Q @ R[:,k]
        R[k,k] = np.linalg.norm(q)
        Q[:,k] = q/R[k,k]

    return Q, R

def eig(A, m = 50, qr="own"):
    """
    Computes the eigenvalues of a square matrix via QR eigenvalue algorithm

    Parameters
    ----------
    A : (nxn) matrix with *distinct* eigenvalues
    m : iteration number
    qr : optional parameter to switch between own qr and scipy qr

    Returns
```

```
    -------
    d : diagonal of the last QR-iterate
    """
    if qr == "own":
        qr = qr_factor
    else:
        from scipy.linalg import qr

    for k in range(m):
        Q, R = qr(A)
        A = R @ Q
    return A.diagonal()


def A_gen(n):
    import numpy as np
    from scipy.linalg import qr
    A = np.random.rand(n,n)
    Q, R = qr(A)
    Lambda = np.diag(np.arange(1,n+1))
    A = Q @ (Lambda @ Q.T)
    return A


if __name__ == "__main__":

  # 2 test
    n = 50
    A = A_gen(n)
    for m in [10, 50, 75, 150]:
        print("number of iterations:\n m =", m, "\napproximate eigenvalues:\n",
              np.sort(np.round(eig(A, m, qr=""), 6)), "\n")
  # 3 compare to numpy.linalg
    print("--> compare to numpy.linalg:\n",np.sort(np.linalg.eigvals(A)))
```

Total Number of Points = 30 (T:24, P:6)