TMA4145 Linear Methods
Fall 2018

Norwegian University of Science and Technology
Department of Mathematical Sciences

Exercise set 12: Solutions

Please justify your answers! The most important part is *how* you arrive at an answer, not the answer itself.

$\boxed{1}$ Let

$$z_1 = \sqrt{\frac{2}{3}}\begin{bmatrix}1\\0\end{bmatrix}, \quad z_2 = \sqrt{\frac{2}{3}}\begin{bmatrix}-1/2\\ \sqrt{3}/2\end{bmatrix}, \quad z_3 = \sqrt{\frac{2}{3}}\begin{bmatrix}-1/2\\ -\sqrt{3}/2\end{bmatrix}.$$

Show that for every $x \in \mathbb{R}^2$ we have

a)

$$\|x\|^2 = \sum_{i=1}^{3} |\langle x, z_i\rangle|^2$$

b)

$$x = \sum_{i=1}^{3} \langle x, z_i\rangle z_i$$

Remark. The vectors z_1, z_2, z_3 span $\mathbb{R}^2$, but they are obviously not an orthonormal basis (they are not even linearly independent). Still, they satisfy a generalization of Parseval's identity and "act like" an orthonormal basis. Such systems appear very naturally in applications (e.g. in signal analysis), and are often called Parseval frames.

Solution. a) Let $x = \begin{bmatrix}x_1\\x_2\end{bmatrix}$ be an arbitrary element of $\mathbb{R}^2$.

$$\begin{aligned}
\sum_{i=1}^{3} |\langle x, z_i\rangle|^2 &= \frac{2}{3}\left(x_1^2 + \left(-\frac{x_1}{2} + \frac{\sqrt{3}x_2}{2}\right)^2 + \left(-\frac{x_1}{2} - \frac{\sqrt{3}x_2}{2}\right)^2\right) \\
&= \frac{2}{3}\left(x_1^2 + \frac{x_1^2}{4} + \frac{3x_2^2}{4} - \frac{\sqrt{3}x_1x_2}{2} + \frac{x_1^2}{4} + \frac{3x_2^2}{4} + \frac{\sqrt{3}x_1x_2}{2}\right) \\
&= \frac{2}{3}\left(\frac{3}{2}x_1^2 + \frac{3}{2}x_2^2\right) \\
&= x_1^2 + x_2^2 = \|x\|^2.
\end{aligned}$$

b) We could show this by direct computation, but here we will use a different approach that is more "informative". It turns out that the Parseval identity in a) implies the representation in b). Indeed, for all $x, y \in \mathbb{R}^2$ we have, by a) and the polarization identity

$$
\begin{aligned}
\langle x, y\rangle &= \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right) \\
&= \frac{1}{4}\sum_{i=1}^{3}\left(\langle x+y, z_i\rangle^2 - \langle x-y, z_i\rangle^2\right) \\
&= \frac{1}{4}\sum_{i=1}^{3}\left((\langle x, z_i\rangle + \langle y, z_i\rangle)^2 - (\langle x, z_i\rangle - \langle y, z_i\rangle)^2\right).
\end{aligned}
$$

After multiplying out and cancelling terms, we are left with

$$\langle x, y\rangle = \sum_{i=1}^{3}\langle x, z_i\rangle\langle y, z_i\rangle = \langle\sum_{i=1}^{3}\langle x, z_i\rangle z_i, y\rangle$$

for all $x, y \in \mathbb{R}^2$. In particular, this implies that

$$x = \sum_{i=1}^{3}\langle x, z_i\rangle z_i,$$

since the previous calculation shows that $\langle x - \sum_{i=1}^{3}\langle x, z_i\rangle z_i, y\rangle = 0$ for all y, meaning that $x - \sum_{i=1}^{3}\langle x, z_i\rangle z_i \in (\mathbb{R}^2)^{\perp} = \{0\}$ – hence $x - \sum_{i=1}^{3}\langle x, z_i\rangle z_i = 0$ and $x = \sum_{i=1}^{3}\langle x, z_i\rangle z_i$.

$\boxed{2}$ Let $\|\cdot\|_a$ and $\|\cdot\|_b$ be equivalent norms on a vector space X. Show that any set $U \subset X$ is open in $(X, \|\cdot\|_a)$ if and only if it is open in $(X, \|\cdot\|_b)$.

Remark. This is in fact a two-way implication; if any set $U \subset X$ is open in $(X, \|\cdot\|_a)$ if and only if it is open in $(X, \|\cdot\|_b)$, then necessarily the norms $\|\cdot\|_a$ and $\|\cdot\|_b$ are equivalent on X.

Solution. Assume that $\|.\|_a$ and $\|.\|_b$ are equivalent norms. This means that we have constants $C_1, C_2 > 0$ such that

$$C_1\|x\|_a \leq \|x\|_b \leq C_2\|x\|_a$$

for all $x \in X$. Then assume that U is open in $(X, \|.\|_a)$, and pick any $x \in U$. To show that U is open in $(X, \|.\|_b)$, we need to find some open ball

$$B_\epsilon^b(x) = \{y \in X : \|y - x\|_b < \epsilon\}$$

such that $B_\epsilon^b(x) \subset U$. Since U is open in $(X, \|.\|_a)$, there exists some $r > 0$ such that

$$B_r^a(x) = \{y \in X : \|y - x\|_a < r\} \subset U.$$

Since $B_r^a(x) \subset U$ and we want $B_\epsilon^b(x) \subset U$, it will clearly be enough to find $\epsilon > 0$ such that

$$B_\epsilon^b(x) \subset B_r^a(x).$$

The key to finding this ϵ is the inequality

$$\|x\|_a \leq \frac{1}{C_1}\|x\|_b.$$

I claim that if we pick $\epsilon = C_1 r$, then $B^b_\epsilon(x) \subset B^a_r(x) \subset U$.
To prove this, assume that $y \in B^b_\epsilon(x)$. We then find that

$$\begin{aligned}\|x-y\|_a &\leq \frac{1}{C_1}\|x-y\|_b \\ &< \frac{1}{C_1}\epsilon = r,\end{aligned}$$

hence $y \in B^a_r(x)$. The proof that any set U that is open in $(X, \|.\|_a)$ whenever U is open in $(X, \|.\|_b)$ is proved the same way, just using the inequality

$$\|x\|_b \leq C_2\|x\|_a.$$

$\boxed{3}$ Suppose that $v_1, \ldots, v_k$ are non-zero eigenvectors of an operator T corresponding to distinct eigenvectors $\lambda_1, \ldots, \lambda_k$. Show that $\{v_1, \ldots v_k\}$ is a linearly independent set.

Solution. We will prove the statement by induction. If $k = 1$ we only have one vector v_1, and since we assume that v_1 is an eigenvector, we know that v_1 is non-zero. Hence the set $\{v_1\}$ is linearly independent; if

$$c_1 v_1 = 0 \quad \text{for some scalar } c_1,$$

then $c_1 = 0$, since if $c_1 \neq 0$ we would have

$$v_1 = \frac{1}{c_1}(c_1 v_1) = 0.$$

Then assume that the statement is true for $k-1$ eigenvalues, and assume that

$$c_1 v_1 + c_2 v_2 + \cdots c_k v_k = 0 \tag{1}$$

where c_j are scalars v_j and non-zero eigenvectors of an operator T corresponding to distinct eigenvectors λ_j. We need to show that $0 = c_1 = c_2 \cdots = c_k$. Since $Tv_k = \lambda_k v_k$, we have that $(T - \lambda_k I)v_k = 0$, where I is the identity operator. We can therefore get rid of v_k in equation (1) by applying $T - \lambda_k I$ to the equation. We then get

$$0 = (T-\lambda_k I)(c_1v_1+c_2v_2+\cdots c_kv_k) = c_1(\lambda_1-\lambda_k)v_1+c_2(\lambda_2-\lambda_k)v_2+\cdots+c_{k-1}(\lambda_{k-1}-\lambda_k)v_{k-1},$$

where we have used that the v_j are eigenvectors. By the induction assumption, $\{v_1, v_2, ..., v_{k-1}\}$ is a linearly independent set. Therefore

$$0 = c_1(\lambda_1 - \lambda_k) = c_2(\lambda_2 - \lambda_k) = ... = c_{k-1}(\lambda_{k-1} - \lambda_k),$$

and since we assume that the eigenvalues are distinct this implies that

$$0 = c_1 = c_2 = ... = c_{k-1}.$$

Therefore equation (1) reads

$$c_k v_k = 0,$$

and as before this implies that also $c_k = 0$, hence $\{v_1, v_2, \ldots, v_k\}$ is a linearly independent set.

$\boxed{4}$ Let T be the shift operator on ℓ^2 defined by $T(x_1, x_2, ...) = (0, x_1, x_2, ...)$.

1. Show that T has no eigenvalues.
2. Does T^* have any eigenvalues?

1. Assume that λ is an eigenvalue of T with eigenvector $y = (y_1, y_2, \ldots)$. Then $Ty = \lambda y$, and writing out both sides we find

$$(0, y_1, y_2, \ldots) = (\lambda y_1, \lambda y_2, \lambda y_3 \ldots). \tag{2}$$

In particular $\lambda y_1 = 0$, which implies that either $y_1 = 0$ or $\lambda = 0$. If $\lambda = 0$, then equality (2) becomes

$$(0, y_1, y_2, \ldots) = (0, 0, \ldots)$$

which shows that $y = 0$, hence not an eigenvector since eigenvectors are non-zero by definition. We may therefore assume that $y_1 = 0$ and $\lambda \neq 0$. In this case the equality (2) becomes

$$(0, 0, y_2, \ldots) = (0, \lambda y_2, \lambda y_3 \ldots), \tag{3}$$

which implies that $y_2 = 0$. Inserting this back into the equation, we find

$$(0, 0, 0, \ldots) = (0, 0, \lambda y_3 \ldots), \tag{4}$$

hence $y_3 = 0$. We may clearly continue like this to show that y is 0 in all coordinates, hence $y = 0$ and y is not an eigenvector.

2. We know from the lectures (and it is not difficult to show) that $T^*(x_1, x_2, \ldots) = (x_2, x_3, \ldots)$. This operator has eigenvalues. For instance, let $y = (1, \frac{1}{2}, \frac{1}{2^2}, \ldots, \frac{1}{2^{p-1}}, \ldots)$. Then

$$T^* y = (\frac{1}{2}, \frac{1}{2^2}, \ldots, \frac{1}{2^p}, \ldots) = \frac{1}{2} y,$$

hence y is an eigenvector with eigenvalue $\frac{1}{2}$. Note that $y \in \ell^2$, which we needed since we defined T on ℓ^2.

$\boxed{5}$ Let U be a $n \times n$ matrix with columns $u_1, ..., u_n$. Show that the following statements are equivalent:

1. U is unitary.
2. $\{u_1, ..., u_n\}$ is an orthonormal basis of $\mathbb{C}^n$.

Solution. Note that the column u_i is of the form $u_i = (u_{1,i}, u_{2,i}, \dots, u_{n,i})^T$, and the inner product between two columns is given by

$$\begin{aligned}\langle u_i, u_j \rangle &= \sum_{k=1}^{n} u_{k,i}\overline{u_{k,j}} \\ &= u_i \cdot u_j^*,\end{aligned}$$

where $u_i \cdot u_j^*$ is the usual dot product for vectors in $\mathbb{C}^n$ and

$$u_j^* = \begin{pmatrix} \overline{u_{1,j}} \\ \overline{u_{2,j}} \\ \dots \\ \overline{u_{n,j}} \end{pmatrix}.$$

<u>$1 \Rightarrow 2$</u>

Assume that $U^*U = I$. Recall that element (i, j) of the matrix product U^*U is the dot product of row i of U^* with column j of U. Since row i of U^* is u_i^*, this means that element (i, j) of U^*U is $u_i^* \cdot u_j$. Furthermore $U^*U = I$, which implies that $u_i^* u_j = \delta_{i,j}$ for $i, j = 1, \dots, n$. Then we have

$$\langle u_j, u_i \rangle = u_i^* \cdot u_j = \delta_{i,j},$$

hence $(u_1, u_2, \dots, u_n)$ is an orthonormal system of vectors in $\mathbb{C}^n$. To show that it is a basis for $\mathbb{C}^n$ it is enough to note that $\mathbb{C}^n$ has dimension n, and the system consists of n vectors. Hence the columns form a linearly independent subset of n vectors in an n-dimensional space, and it follows that the columns form a basis.

<u>$2 \Rightarrow 1$</u>

Assume that the columns $u_1, u_2, \dots, u_n$ of U are an orthonormal basis of $\mathbb{C}^n$, i.e.

$$\langle u_i, u_j \rangle = \delta_{i,j},$$

for $i, j = 1, \dots, n$. Then we have

$$u_i \cdot u_j^* = \langle u_i, u_j \rangle = \delta_{i,j},$$

hence we have $U^*U = I$. One gets that $U^*U = I$ from exactly the same argument, or by knowing that a left inverse of matrix is also a two-sided inverse.

$\boxed{6}$ Given the matrix

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 2 \\ 2 & 1 \end{pmatrix}.$$

a) Compute the singular value decomposition of A.

b) Use the result of a) to find:

1. The pseudo-inverse of A.
2. Find the minimal norm solution of $Ax = b$ for $b = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$.

Solution. a) We need to find matrices U, V, Σ with certain properties such that $A = U\Sigma V^*$. If we look back to the proof of the SVD in the lecture notes (theorem 7.20), we can deduce how we construct U, V and Σ. Note that since A is a 3×2-matrix with rank 2, we will find that V is a 2×2 matrix, Σ a 3×2-matrix and U a 3×3-matrix.

1. V is picked as the matrix that diagonalizes A^*A, and therefore the columns of U are the normalized eigenvectors of A^*A.
2. Σ is the 3×2 matrix with the positive singular values of A (= the square roots of the positive eigenvalues of A^*A) along the diagonal, and zero elsewhere. [1]
3. The first 2 columns of the 3×3 U will be the vectors Av_1 and Av_2, where v_1, v_2 are the columns of V. Since we need U to be a 3×3 matrix, we need one more column v_3, which we find by picking a normalized vector v_3 that is orthogonal to v_1 and v_2.

Let us now find these matrices.

1. $A^*A = \begin{pmatrix} 9 & 8 \\ 8 & 9 \end{pmatrix}$ has as characteristic polynomial $x^2 - 18x + 17 = (x - 17)(x - 1)$. Hence the eigenvalues of A^*A are $\lambda_1 = 17$ and $\lambda_2 = 1$. The corresponding normalized eigenvectors are $v_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $v_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Consequently, we have
$$V = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}.$$

[1] If we did not have enough singular values to fill the diagonal of Σ, we would just put zeros in the rest of the diagonal of Σ. This will not be an issue in this example.

2. The singular values of A are $\sigma_1 = \sqrt{17}$ and $\sigma_2 = 1$. Thus
$$\Sigma = \begin{pmatrix} \sqrt{17} & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

3. The first two columns of U are given by
$$u_1 = \frac{1}{\sqrt{17}}\frac{1}{\sqrt{2}}Av_1 = \frac{1}{\sqrt{34}}\begin{pmatrix} 1 & 2 \\ 2 & 2 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{34}}\begin{pmatrix} 3 \\ 4 \\ 3 \end{pmatrix}$$
and by
$$u_2 = \frac{1}{1}Av_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 2 \\ 2 & 2 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}.$$
Consequently, U has the form
$$\begin{pmatrix} \frac{3}{\sqrt{34}} & \frac{-1}{\sqrt{2}} & x_1 \\ \frac{4}{\sqrt{34}} & 0 & x_2 \\ \frac{3}{\sqrt{34}} & \frac{1}{\sqrt{2}} & x_3 \end{pmatrix}$$
The last column is determined by the assumption that it has to be orthogonal to the first two columns. The choice
$$u_3 = \frac{1}{\sqrt{17}}\begin{pmatrix} 2 \\ -3 \\ 2 \end{pmatrix}$$
satisfies these conditions, but there are many other ways to complete the first two columns to become an orthonormal basis for $\mathbb{C}^3$,
$$\begin{pmatrix} \frac{3}{\sqrt{34}} & \frac{-1}{\sqrt{2}} & \frac{2}{\sqrt{17}} \\ \frac{4}{\sqrt{34}} & 0 & \frac{-3}{\sqrt{17}} \\ \frac{3}{\sqrt{34}} & \frac{1}{\sqrt{2}} & \frac{2}{\sqrt{17}} \end{pmatrix}.$$

4. The SVD of A is
$$\begin{pmatrix} \frac{3}{\sqrt{34}} & \frac{-1}{\sqrt{2}} & \frac{2}{\sqrt{17}} \\ \frac{4}{\sqrt{34}} & 0 & \frac{-3}{\sqrt{17}} \\ \frac{3}{\sqrt{34}} & \frac{1}{\sqrt{2}} & \frac{2}{\sqrt{17}} \end{pmatrix}\begin{pmatrix} \sqrt{17} & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}.$$

b)

1. By theorem 7.23 in the notes, $A^\dagger = V\Sigma^+U^*$ where Σ^+ is the matrix obtained from Σ by replacing the singular values σ_i with σ_i^{-1} and taking the transpose. Hence
$$A^\dagger = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}\begin{pmatrix} \frac{1}{\sqrt{17}} & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} \frac{3}{\sqrt{34}} & \frac{4}{\sqrt{34}} & \frac{3}{\sqrt{34}} \\ \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{2}{\sqrt{17}} & \frac{-3}{\sqrt{17}} & \frac{2}{\sqrt{17}} \end{pmatrix} = \begin{pmatrix} \frac{-7}{17} & \frac{2}{17} & \frac{10}{17} \\ \frac{10}{17} & \frac{2}{17} & \frac{-7}{17} \end{pmatrix}.$$

2. By theorem 7.25 the notes, the least squares solution is given by

$$A^\dagger b = \begin{pmatrix} \frac{-7}{17} & \frac{2}{17} & \frac{10}{17} \\ \frac{10}{17} & \frac{2}{17} & \frac{-7}{17} \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} \frac{27}{17} \\ \frac{-7}{17} \end{pmatrix}.$$