EECS 127/227AT **Fall 2020**

Solutions for Discussion 2

1 Symmetric Matrices and Quadratic Forms

(a) A general quadratic function $f : \mathbb{R}^n \to \mathbb{R}$ can be written as

$$f(x) = x^T Ax + x^T b + c,$$

where $A \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^n$ and $c \in \mathbb{R}$. Note that $x^T Ax$ is a number, so if we take the transpose we get

$$x^T Ax = x^T A^T x,$$

from which we get

$$x^T Ax = x^T \left(\frac{A + A^T}{2}\right) x,$$

so, when discussing a quadratic function, there is no loss of generality in assuming that A is symmetric. In the special case where $b = 0$ and $c = 0$, and assuming A is symmetric, we get the quadratic function $x^T Ax$ on $\mathbb{R}^n$. Such a function is called a *quadratic form*.

Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix, i.e. $A = A^T$. In class we have shown that the eigenvalues of A are real and that after an orthogonal change of basis A can be expressed as a diagonal matrix, i.e. we can write $A = U^T \Lambda U$ where U is an orthogonal matrix and Λ is a diagonal matrix.

In this problem we will understand the quadratic form associated to a symmetric positive semidefinite matrix A by considering the set

$$\mathcal{E} := \{x \in \mathbb{R}^n : x^T Ax \le 1\}.$$

Note that if A were the identity matrix then $\mathcal{E}$ would be the unit ball in $\mathbb{R}^n$, so in general $\mathcal{E}$, as defined above, can be considered the analog of the unit ball for the quadratic form defined by $A \in \mathbb{S}^n_+$.

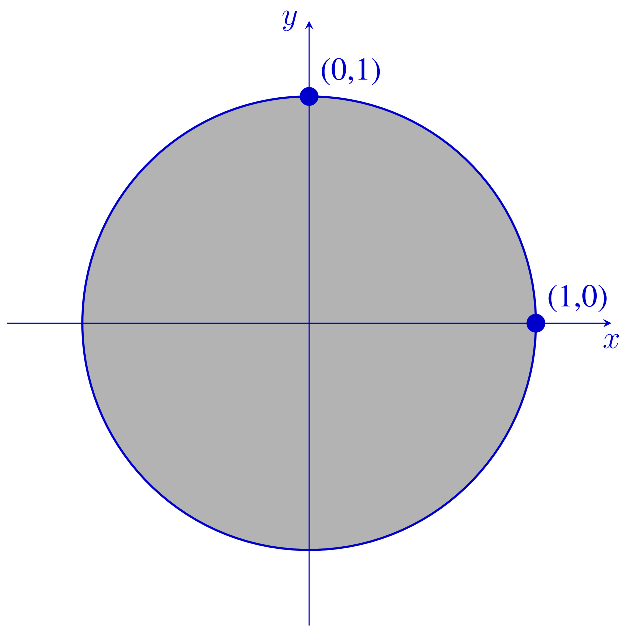
Sketch the set $\mathcal{E}$ for the following matrices and think of how the eigenvalues and eigenvectors of A relate to the shape of the set $\mathcal{E}$.

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

Note that here it is obvious that A is symmetric and positive semidefinite.

Solution:

Notice that $x^T Ax = x_1^2 + x_2^2$, therefore $\mathcal{E}$ is simply the region inside the unit circle in $\mathbb{R}^2$, including the boundary, i.e. the closed unit disk in $\mathbb{R}^2$.

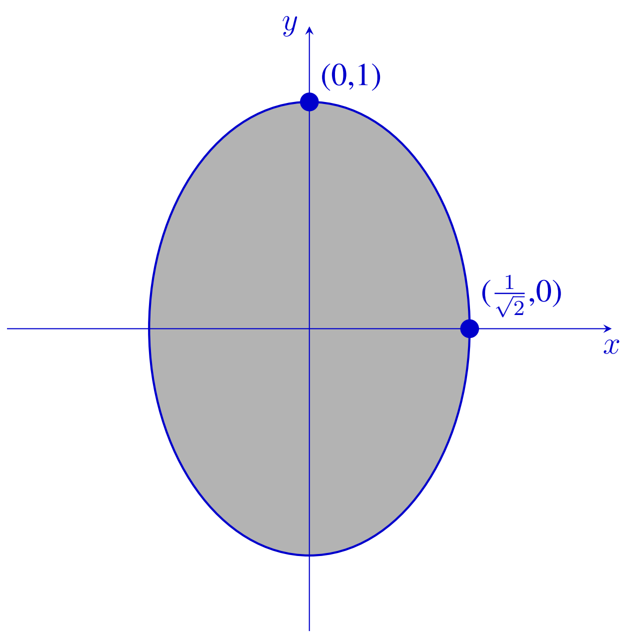


ii. $A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$

Note that here it is still obvious that A is symmetric and positive semidefinite.

Solution:

Expanding out the quadratic form as we did in part (i), we see that $\mathcal{E} = \{x \in \mathbb{R}^2 : 2x_1^2 + x^2 \leq 1$. This is the closed region bounded by an ellipse, as shown in the figure.



iii. $A = \begin{bmatrix} 1.5 & 0.5 \\ 0.5 & 1.5 \end{bmatrix}$

Here A is symmetric but it is not immediately obvious that it positive semidefinite. For this we need to compute the eigenvalues. Here this is easy to do by computing the trace (which is the sum of the eigenvalues) and the determinant (which is the product of the eigenvalues). From this we can conclude that A is positive semidefinite, since the trace is 3 and the determinant is 2, so the eigenvalues must be 2 and 1.

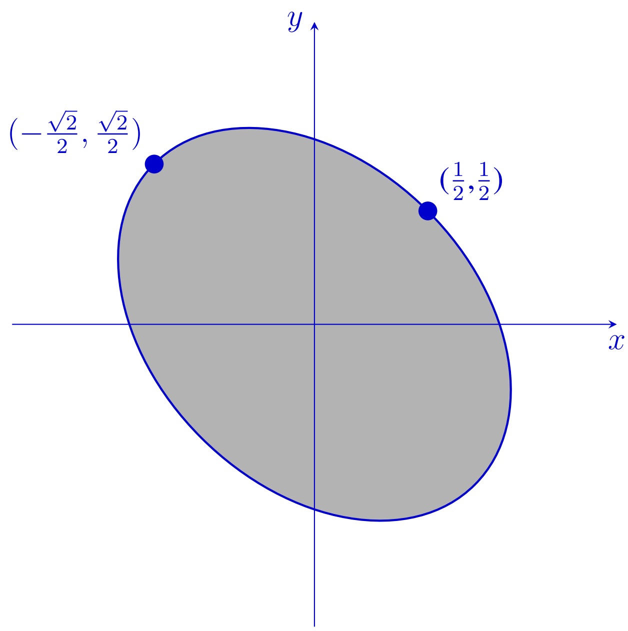
Solution:
Because A is symmetric, we know that we can write $A = U\Lambda U^T$, where U is orthogonal and Λ is diagonal. This is useful when we consider the quadratic form $x^T Ax$, because we can write

$$\begin{aligned} x^T Ax &= x^T U\Lambda U^T x && (1)\\ &= x^T U\Lambda^{\frac{1}{2}}\Lambda^{\frac{1}{2}}U^T x && (2)\\ &= \|\Lambda^{\frac{1}{2}}U^T x\|_2^2 && (3)\\ &= \|\Lambda^{\frac{1}{2}}\tilde{x}\|_2^2, && (4) \end{aligned}$$

where $\tilde{x} := U^T x$. Thus, if Λ has the diagonal entries λ_1 and λ_2, the set $\mathcal{E}$ can be expressed as

$$\mathcal{E} = \{x \in \mathbb{R}^2 \ : \ \lambda_1\tilde{x}_1^2 + \lambda_2\tilde{x}_2^2 \leq 1\}.$$

Here we find $U = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$, and $\Lambda = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$.
You should check that U, as given, is an orthogonal matrix. In fact, U is the rotation matrix for the angle $\theta = \frac{\pi}{4}$. Comparing the expression for $\mathcal{E}$ in the $\tilde{x}$ coordinates to the set $\mathcal{E}$ in part (ii) of the problem, since $\lambda_1 = 2$ and $\lambda_2 = 1$, we see that $\mathcal{E}$ in this problem can be obtained by rotating the ellipse we obtained in part (ii) by $\frac{\pi}{4}$ counter clockwise. The result is drawn in the following figure.



(b) Let $A \in \mathbb{S}_+^n$ (i.e. A is a symmetric positive semidefinite matrix), $b \in \mathbb{R}^n$ and $c \in \mathbb{R}$. Consider the following minimization problem:

$$p^* = \min_{x\in\mathbb{R}^n} x^T Ax + b^T x + c.$$

We will now study this problem in some examples.

Find the optimal value of this optimization problem, i.e. p^*, for the following cases:

(i) $A = \begin{bmatrix} \frac{3}{4} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{3}{4} \end{bmatrix}$, $b = 0$, $c = 0$.
You should check that here we have $A \in \mathbb{S}^n_+$.

Solution:
Since $A \in \mathbb{S}^n_+$, the objective function, which is $x^T Ax$ since $b = 0$, satisfies $x^T Ax \geq 0$ for all x. At $x = 0$ the value of the objective function is 0. Therefore $p^* = 0$. In fact, you can check that $A \in \mathbb{S}^n_{++}$ (i.e. it is symmetric, positive semidefinite and invertible) because the eigenvalues of A are $\lambda_1 = \frac{1}{2}$ and $\lambda_2 = 1$. Therefore the optimal value p^* is achieved precisely at $x = 0$—no other setting for x achieves the optimum.

(ii) $A = \begin{bmatrix} \frac{3}{4} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{3}{4} \end{bmatrix}$, $b = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$, $c = 0$.

Solution:
Let us write $A = U^T \Lambda U$, where U is orthogonal and Λ is diagonal, with the diagonal entries $\lambda_1 = \frac{1}{2}$ and $\lambda_2 = 1$. The objective function of the optimization problem can then be written as

$$\begin{aligned} x^T Ax + b^T x &= x^T (U^T \Lambda U)x + b^T U^T U x \\ &= (Ux)^T \Lambda (Ux) + (Ub)^T (Ux) \\ &= \tilde{x}^T \Lambda \tilde{x} + \tilde{b}^T \tilde{x}, \end{aligned}$$

where we define $\tilde{x} := Ux$ and $\tilde{b} := Ub$. Thus the objective function is

$$\sum_{i=1}^{2} \left(\lambda_i \tilde{x}_i^2 + \tilde{b}_i \tilde{x}_i \right).$$

This is the sum of two objectives, one for each coordinate, and so minimizing the sum is equivalent to minimizing each expression individually. We see that the minimum occurs at x^* where $\tilde{x}_i^* = -\frac{\tilde{b}_i}{2\lambda_i}$ for $i = 1, 2$. In vector form, this reads

$$\tilde{x}^* = -\frac{1}{2} \Lambda^{-1} \tilde{b},$$

which in the original coordinates reads

$$x^* = -\frac{1}{2} U^T \Lambda^{-1} U b,$$

(by multiplying both sides on the left by U^T). But we have $A^{-1} = U^T \Lambda^{-1} U$. Therefore the objective is minimized at

$$x^* = -\frac{1}{2} A^{-1} b. \tag{5}$$

Substituting x^* for x in the objective function gives

$$p^* = -\frac{1}{4} b^T A^{-1} b.$$

In this example, we have $A^{-1} = \begin{bmatrix} 1.5 & 0.5 \\ 0.5 & 1.5 \end{bmatrix}$. Therefore the optimum is uniquely achieved at $x^* = -\frac{1}{2} A^{-1} b = [-\frac{1}{2} \quad -\frac{1}{2}]^T$. and the optimal value is $p^* = -0.25$.

It is also instructive to derive the formula in equation (5) for the optimizer in part (b)(ii) of this problem using the notation of multivariable calculus. Consider a function

$$f : \mathbb{R}^n \to \mathbb{R}$$

Assuming the function is differentiable at $x \in \mathbb{R}^n$, its gradient at x is given by

$$\nabla f(x) := \begin{bmatrix} \frac{\partial f}{\partial x_1}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{bmatrix}.$$

The i-th coordinate of $\nabla f(x)$ can be written as $e_i^T \nabla f(x)$, where e_i denotes the standard unit vector in direction i. From the definition of the derivative, we have

$$e_i^T \nabla f(x) = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \left(f(x + \epsilon e_i) - f(x) \right).$$

Consider now the quadratic function $f : \mathbb{R}^n \to \mathbb{R}$ given by

$$f(x) := x^T A x + b^T x + c, \tag{6}$$

where $A \in \mathbb{S}^n$, $b \in \mathbb{R}^n$ and $c \in \mathbb{R}$. We can then write

$$\begin{aligned} e_i^T \nabla f(x) &= \lim_{\epsilon \to 0} \frac{1}{\epsilon} \left(\left((x + \epsilon e_i)^T A (x + \epsilon e_i) + b^T (x + \epsilon e_i) + c \right) - \left(x^T A x + b^T x + c \right) \right) \\ &= e_i^T A x + x^T A e_i + b^T e_i \\ &= e_i^T \left((A + A^T) x + b \right) \\ &= e_i^T \left(2Ax + b \right). \end{aligned}$$

Since this holds for all $1 \leq i \leq n$, we have

$$\nabla f(x) = 2Ax + b.$$

The condition for x to be an extremum of the quadratic function in equation (6) when A is symmetric becomes

$$2Ax + b = 0.$$

Here we note that if A is invertible then there is a unique extremum, given by

$$x^* = -2A^{-1} b,$$

while if A is not invertible the set of extrema will be an affine translate of the null space of A. Returning to the case of a general $f : \mathbb{R}^n \to \mathbb{R}$, to check if an extremum x is a local maximum, a local minimum, or a saddle point, we need to compute the second order derivative, i.e. the *Hessian* of f at x. This is the matrix of second order partial derivatives of f at x, given by

$$\nabla^2 f(x) := \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix} (x).$$

The Hessian is the *Jacobian* of the gradient, when the gradient is viewed as a function from $\mathbb{R}^n$ to $\mathbb{R}^n$ on its domain of definition. To compute the ij-entry of the Hessian therefore, we need to compute the derivative of the j-th component of the gradient in the direction i (or we could compute the derivative of the i-th component of the gradient in the direction j, which is the same thing). In the notation of multivariable calculus, this can be written as

$$e_i^T \nabla^2 f(x) e_j = e_i^T \nabla (e_j^T \nabla f)(x) = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \left(e_j^T \nabla f(x + \epsilon e_i) - e_j^T \nabla f(x) \right).$$

Returning to the quadratic function of equation (6) we have already computed $\nabla f(x) = (A + A^T)x + b$. Hence we get

$$\begin{aligned} e_i^T \nabla^2 f(x) e_j &= \lim_{\epsilon \to 0} \frac{1}{\epsilon} \left(e_j^T \left((A + A^T)(x + \epsilon e_i) + b \right) - e_j^T \left((A + A^T)x + b \right) \right) \\ &= e_j^T (A + A^T) e_i \\ &= e_i^T (A + A^T) e_j, \end{aligned}$$

which tells us that the Hessian is given by

$$\nabla^2 f(x) = A + A^T.$$

When A is a symmetric matrix, this becomes

$$\nabla^2 f(x) = 2A.$$

Further, when $A \in \mathbb{S}_+^n$, i.e. A is a symmetric positive semidefinite matrix, we will have

$$z^T \nabla^2 f(x) z = 2z^T A z \geq 0,$$

for all $z \in \mathbb{R}^n$, so every extremum will be a local minimum.

In particular, since there was a unique extremum in the case where A was symmetric and invertible, if A is also positive semidefinite then this unique extremum, i.e. $x^* = -2A^{-1}b$ is the global minimum of the quadratic function in equation (6), as we saw in the example in part (b) (ii) of this problem.

Remark: It is conventional in optimization theory to restrict the use of the terminology "positive semidefinite" to symmetric matrices, so when we say a matrix is positive semidefinite we automatically imply that it is symmetric. We will nevertheless often use the terminology "symmetric and postive semidefinite", as we have done in this problem, in order to emphasize the requirement of symmetry. Note that to check that $A \in \mathbb{R}^{n \times n}$ is positive semidefinite, as the terminology is defined in optimization theory, you not only have to check that $x^T A x \geq 0$ for all $x \in \mathbb{R}^n$, but you also have to check that A is symmetric.

2 Cauchy matrix

(a) Let $v \in \mathbb{R}^n$. Let $A := vv^T$. Note that $A \in \mathbb{R}^{n\times n}$.

Recall that $\mathbb{S}^n$ is our notation for the set of symmetric matrices in $\mathbb{R}^{n\times n}$, and that $\mathbb{S}^n_+$ is our notation for the subset of $\mathbb{S}^n$ comprising symmetric positive semidefinite matrices.

Show that $A \in \mathbb{S}^n_+$. What is the rank of A?

Solution:
Since

$$A^T = (vv^T)^T = v^{TT}v^T = vv^T = A,$$

we see that A is symmetric, i.e. $A \in \mathbb{S}^n$.

Further, for every $x \in \mathbb{R}^n$ we have

$$x^T Ax = x^T(vv^T)x = (x^Tv)(v^Tx) = (x^Tv)^2 \geq 0.$$

Since this holds for every $x \in \mathbb{R}^n$, we see that A is positive semidefinite, i.e. $A \in \mathbb{S}^n_+$.

A is a dyad. If $v = 0$, the rank of A is zero. If v is nonzero the rank of A is 1.

(b) Let $A, B \in \mathbb{S}^n_+$. Show that $A + B \in \mathbb{S}^n_+$.

Solution:
To check that $A + B$ is a symmetric matrix, write

$$(A + B)^T = A^T + B^T \overset{(*)}{=} A + B,$$

where in step (*) we have used the fact that A and B are symmetric matrices. To check that $A + B$ is positive semidefinite, given any $x \in \mathbb{R}^n$, write

$$x^T(A + B)x = x^T Ax + x^T Bx \overset{(**)}{\geq} 0,$$

where in step (**) we have used the fact that A and B are positive semidefinite matrices. Since this holds for all $x \in \mathbb{R}^n$, we conclude that $A + B \in \mathbb{S}^n_+$.

(c) Let $(A_k \in \mathbb{S}^n_+)_{k\geq 1}$ be a sequence of symmetric positive semidefinite $n \times n$ matrices. Suppose that $\lim_{k\to\infty} A_k = A$ in the sense that the individual entries $(A_k)_{ij}$ converge to $(A)_{ij}$ as $k \to \infty$ for $1 \leq i, j \leq n$.

Show that $A \in \mathbb{S}^n_+$.

Solution:
Since $(A_k)_{ij} = (A_k)_{ji}$ for all $k \geq 1$ and all $1 \leq i, j \leq n$, by taking limits as $k \to \infty$ we have $(A)_{ij} = (A)_{ji}$ for all $1 \leq i, j \leq n$. This shows that A is a symmetric matrix.

For every $x \in \mathbb{R}^n$, we have

$$x^T A_k x = \sum_{i=1}^n \sum_{j=1}^n (A_k)_{ij} x_i x_j \geq 0.$$

Taking the limit as $k \to \infty$, we get

$$x^T Ax = x^T \left(\lim_{k\to\infty} A_k\right) x = \sum_{i=1}^n \sum_{j=1}^n \left(\lim_{k\to\infty} (A_k)_{ij}\right) x_i x_j = \lim_{k\to\infty} \sum_{i=1}^n \sum_{j=1}^n (A_k)_{ij} x_i x_j \geq 0.$$

So we get

$$x^T A x \geq 0.$$

Since this holds for all $x \in \mathbb{R}^n$, we conclude that $A \in \mathbb{S}^n_+$.

(d) Let $\alpha_1, \ldots, \alpha_n$ be n strictly positive real numbers. The associated *Cauchy matrix* is the matrix $A \in \mathbb{R}^{n \times n}$ whose (i, j) entry is $\frac{1}{\alpha_i + \alpha_j}$, i.e.

$$A := \begin{bmatrix} & \vdots & \\ \cdots & \frac{1}{\alpha_i+\alpha_j} & \cdots \\ & \vdots & \end{bmatrix} \in \mathbb{R}^{n \times n}.$$

Show that $A \in \mathbb{S}^n_+$.

Hint:

$$\frac{1}{\alpha_i + \alpha_j} = \int_0^\infty e^{-(\alpha_i+\alpha_j)t} dt.$$

Solution:

For $t \geq 0$, let $v(t) \in \mathbb{R}^n$ be defined via

$$v(t) := \begin{bmatrix} e^{-\alpha_1 t} & \ldots & e^{-\alpha_n t} \end{bmatrix}^T.$$

Using the hint, we see that

$$A = \int_0^\infty v(t)v(t)^T dt.$$

For each $t \geq 0$ the matrix $v(t)v(t)^T \in \mathbb{R}^{n \times n}$ is a symmetric and positive semidefinite matrix, as was seen in part (a) of the problem. Part (b) of the problem tells us that $\mathbb{S}^n_+$ is closed under addition. Part (c) of the problem tells us that $\mathbb{S}^n_+$ is a closed set, i.e. any limits of a sequence in $\mathbb{S}^n_+$ lies in $\mathbb{S}^n_+$. Since an integral is just a limit of sums, we should now expect that A would be in $\mathbb{S}^n_+$.

This can be verified directly. First of all, it can be seen directly from its definition that A is symmetric. Also, for every $x \in \mathbb{R}^n$, we have

$$x^T A x = x^T \left(\int_0^\infty v(t)v(t)^T dt \right) x = \int_0^\infty x^T v(t)v(t)^T x dt = \int_0^\infty (x^T v(t))^2 dt \geq 0.$$

Since this holds for all $x \in \mathbb{R}^n$, we conclude that $A \in \mathbb{S}^n_+$.

3 Adding a dyad

(a) Let $A \in \mathbb{S}^n$ (note that positive semidefiniteness is not assumed in this part of the problem).

Show that $\mathcal{N}(A)$ and $\mathcal{R}(A)$ are orthogonal complements of each other.

Solution:
This is immediate from the fundamental theorem of linear algebra, which tells us that for any $A \in \mathbb{R}^{m\times n}$ the subspaces $\mathcal{N}(A^T)$ and $\mathcal{R}(A)$ of $\mathbb{R}^m$ are orthogonal complements of each other. All we need to observe is that $A^T = A$ (and also $m = n$).

(b) Let $A \in \mathbb{S}^n_+$ and $x \in \mathbb{R}^n$.

Show that

$$x \in \mathcal{N}(A) \Leftrightarrow x^T A x = 0.$$

Solution:
If $x \in \mathcal{N}(A)$ then $Ax = 0$ (by the definition of $\mathcal{N}(A)$) and so $x^T Ax = 0$. Therefore one direction of the claim is obviously true, and in fact is true for any square matrix.

For the other direction, we will appeal to our knowledge that a symmetric matrix A can be diagonalized by an orthogonal change of basis and that the resulting diagonal matrix has nonnegative diagonal entries when the symmetric matrix A is positive semidefinite.

Following this approach, we write $A = U^T \Lambda U$ where U is an orthogonal matrix and Λ is a diagonal matrix with nonnegative entries, and get

$$x^T Ax = 0 \Leftrightarrow x^T U^T \Lambda U x = 0 \Leftrightarrow (\Lambda^{\frac{1}{2}} Ux)^2 = 0 \Leftrightarrow \Lambda^{\frac{1}{2}} Ux = 0.$$

Here $\Lambda^{\frac{1}{2}} \in \mathbb{R}^{n\times n}$ is a square root of Λ, i.e. a diagonal matrix whose square equals Λ. Such a matrix exists in $\mathbb{R}^{n\times n}$ because the diagonal matrix Λ has nonnegative diagonal entries (simply take a real square root – the sign does not matter for our purposes – of each of the diagonal entries of Λ one by one).

But we also have

$$\Lambda^{\frac{1}{2}} Ux = 0 \Longrightarrow (U^T \Lambda^{\frac{1}{2}})\Lambda^{\frac{1}{2}} Ux = 0 \Longrightarrow Ax = 0 \Longrightarrow x \in \mathcal{N}(A).$$

This completes the proof of what was claimed.

Note: In the second part of the proof we have used the assumption that A is symmetric. For instance, for the non-symmetric matrix $A := \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ and $x := \begin{bmatrix} 1 & 0 \end{bmatrix}^T$, we have $x^T Ax = 0$, but $x \notin \mathcal{N}(A)$.

(c) Let $v, w \in \mathbb{R}^n$. Show that

$$\mathcal{R}(vv^T) \subseteq \mathcal{R}(vv^T + ww^T).$$

Hint: When v is nonzero, write $w = (w - \frac{w^T v}{\|v\|_2^2} v) + \frac{w^T v}{\|v\|_2^2} v$.

Solution:
If either $v = 0$ or $w = 0$ the claim is obvious, so we may assume that they are both nonzero.

We have $vv^T \in \mathbb{S}^n_+$ and its range space is the span of its columns, which are all proportional to v, so

$$\mathcal{R}(vv^T) = \text{span}(\{v\}) = \{\alpha v \ : \ \alpha \in \mathbb{R}\},$$

and this has dimension 1, since we assumed that v is nonzero. What we need to show therefore is

$$v \overset{?}{\in} \mathcal{R}(vv^T + ww^T).$$

Note the convention of writing an expression (equality, inequality, containment etc.) with a question mark on it to indicate that it is something we want to establish but have not yet established. We will use this convention often in this course.

Using the hint, we can write

$$\begin{aligned} ww^T &= \left((w - \frac{w^Tv}{\|v\|_2^2}v) + \frac{w^Tv}{\|v\|_2^2}v\right)\left((w - \frac{w^Tv}{\|v\|_2^2}v) + \frac{w^Tv}{\|v\|_2^2}v\right)^T \\ &= (w - \frac{w^Tv}{\|v\|_2^2}v)(w - \frac{w^Tv}{\|v\|_2^2}v)^T + \left(\frac{w^Tv}{\|v\|_2^2}\right)^2 vv^T, \end{aligned}$$

where in the second step the cross term vanishes because

$$(w - \frac{w^Tv}{\|v\|_2^2}v)^T v = w^Tv - w^Tv = 0.$$

This gives

$$vv^T + ww^T = \left(1 + \left(\frac{w^Tv}{\|v\|_2^2}\right)^2\right) vv^T + (w - \frac{w^Tv}{\|v\|_2^2}v)(w - \frac{w^Tv}{\|v\|_2^2}v)^T \tag{7}$$

If w were proportional to v we would have $w = \frac{w^Tv}{\|v\|_2^2}v$ and so equation (7) would tell us that $vv^T + ww^T = \left(1 + \left(\frac{w^Tv}{\|v\|_2^2}\right)^2\right) vv^T$. Since $1 + \left(\frac{w^Tv}{\|v\|_2^2}\right)$ is strictly positive, this would imply that $\mathcal{R}(vv^T)$ equals $\mathcal{R}(vv^T + ww^T)$, which certainly proves what is desired (and more). So we may assume that w is not proportional to v.

Now, we have

$$\mathcal{R}(vv^T) = \{\alpha v \; : \; \alpha \in \mathbb{R}\}$$

and we have

$$\alpha v = (vv^T)(\frac{\alpha}{\|v\|_2^2}v).$$

To show that $\alpha v \in \mathcal{R}(vv^T + ww^T)$ we use the equation (7) to write

$$\alpha v = (vv^T + ww^T)\left(\frac{\frac{\alpha}{\|v\|_2^2}}{1 + \left(\frac{w^Tv}{\|v\|_2^2}\right)^2}v\right),$$

where we have used the orthogonality of $w - \frac{w^Tv}{\|v\|_2^2}v$ and v, This completes the proof of the claim.

(d) Let $A \in \mathbb{S}_+^n$ and $v \in \mathbb{R}^n$. From part (b) of preceding problem, we know that $A + vv^T \in \mathbb{S}_+^n$.

What possible values can $\text{rank}(A + vv^T) - \text{rank}(A)$ take? Explain under what scenarios each of these values results.

Solution:
In part (b) of this problem we have shown that, because $A \in \mathbb{S}^n_+$, we have, for any $x \in \mathbb{R}^n$, that

$$x \in \mathcal{N}(A) \Leftrightarrow x^T A x = 0.$$

Since $A + v^T v \in \mathbb{S}^n_+$, we also have

$$x \in \mathcal{N}(A + vv^T) \Leftrightarrow x^T (A + vv^T) x = 0.$$

Since $x^T(A + vv^T)x = x^T A x + (x^T v)^2$, and we have $x^T A x \geq 0$ and $(x^T v)^2 \geq 0$, it follows that

$$x \in \mathcal{N}(A + vv^T) \Longrightarrow x \in \mathcal{N}(A).$$

From part (a) of this problem it then follows that

$$x \in \mathcal{R}(A) \Longrightarrow x \in \mathcal{R}(A + vv^T). \tag{8}$$

We therefore conclude, in particular, that

$$\text{dim}(\mathcal{R}(A + vv^T)) - \text{dim}(\mathcal{R}(A)) \geq 0.$$

Since $x^T(A + vv^T)x = x^T A x + (x^T v)^2$ we can also conclude that $x^T(A + vv^T)x = 0$ iff we have both $x^T A x = 0$ and $x^T v = 0$, which then tells us that

$$\mathcal{N}(A + vv^T) = \mathcal{N}(A) \cap \mathcal{N}(vv^T). \tag{9}$$

Since $\mathcal{N}(vv^T)$ is the orthogonal complement of $\mathcal{R}(vv^T)$ and since $\dim(\mathcal{R}(vv^T)) = 1$, we have $\dim(\mathcal{N}(vv^T)) = n - 1$. From (9), we therefore get

$$\dim(\mathcal{N}(A + vv^T)) \geq \dim(\mathcal{N}(A)) - 1.$$

We can then conclude that

$$\text{dim}(\mathcal{R}(A + vv^T)) - \text{dim}(\mathcal{R}(A)) \leq 1.$$

We have therefore proved that $\dim(\mathcal{R}(A + vv^T)) - \dim(\mathcal{R}(A))$ is either 0 or 1. Also we have seen from the argument above that

$$\begin{aligned}
\dim(\mathcal{R}(A + vv^T)) - \dim(\mathcal{R}(A)) = 0 \quad &\Leftrightarrow \quad \dim(\mathcal{N}(A + vv^T)) = \dim(\mathcal{N}(A)) \\
&\Leftrightarrow \quad \mathcal{N}(A + vv^T) = \mathcal{N}(A) \\
&\Leftrightarrow \quad \forall x \in \mathbb{R}^n \text{ we have the implication } x^T A x = 0 \Longrightarrow x^T v = 0 \\
&\Leftrightarrow \quad v \in \mathcal{R}(A)
\end{aligned}$$

Thus $\dim(\mathcal{R}(A+vv^T)) - \dim(\mathcal{R}(A)) = 0$ iff $v \in \mathcal{R}(A)$ (and $\dim(\mathcal{R}(A+vv^T)) - \dim(\mathcal{R}(A)) = 1$ iff $v \notin \mathcal{R}(A)$). This explains in which scenarios each of the two cases arises.

Note: The statement in equation (8) is quite subtle. Make sure you understand it. The special case that we worked out in part (c) of this problem should help you to understand why the statement in equation (8) is true.