Problem 1
If P is an orthogonal projector, then $I - 2P$ is unitary.

Proof. We may show this directly. Recall P being an orthogonal projector means that P is both idempotent and symmetric. That is, $P^2 = P$ and $P^T = P$. Recall matrix multiplication is distributive. Now observe,

$$\begin{aligned}(I-2P)(I-2P)^* &= (I-2P)(I^* - 2P^*) \\ &= (I-2P)I^* + (I-2P)(-2P^*) \\ &= I - 2P - 2P^* + 4PP^*\end{aligned} \tag{1}$$

As mentioned by Dongwook in class, this orthogonal projector is tacitly assumed to be a real matrix. Then it follows that $P^* = P^T$. Since P is symmetric, $P^* = P^T = P$. Then equation (1) can be rewritten as follows:

$$(I-2P)(I-2P)^* = I - 2P - 2P^* + 4PP^* = I - 4P + 4P^2$$
$$\implies (I-2P)(I-2P)^* = I - 4P + 4P$$
$$\implies (I-2P)(I-2P)^* = I$$

Similarly, we should show $(I-2P)^*(I-2P) = I$. Well

$$\begin{aligned}(I-2P)^*(I-2P) &= (I^* - 2P^*)(I-2P) \\ &= I - 2P - 2P^* + 4P^2 \\ &= I - 4P + 4P \\ &= I\end{aligned} \tag{2}$$

and thus $(I-2P)^*(I-2P) = (I-2P)(I-2P)^* = I$, so by definition $I - 2P$ is unitary as desired. $\square$

Problem 2
Let $P \in \mathbb{R}^{m\times m}$ be a nonzero projector.

(a) Show that $||P||_2 \geq 1$, with equality if and only if P is an orthogonal projector

(b) If P is an orthogonal projector, then P is positive semi-definite and has eigenvalues either zero or one.

Solution:

(a):

Proof. Since P is a nonzero projector, $P^2 = P$. If we assume that additionally P is an orthogonal projector, that is $P = P^T$, then we have the following. By the lecture notes week 3 after line (3.41), a square orthogonal projector can be constructed as

$$P = QQ^T$$

for $Q \in \mathbb{R}^{m\times n}$ with columns that form an orthonormal basis for the span(A) and satisfy $Q^TQ = I$. Then

$$||P||_2 = ||QQ^T||_2 = ||Q||_2\,||Q^T||_2$$

where the last equality is *truly* an equality, as opposed to an inequality, as a result of working with the 2-norm. We can continue to rewrite using this 2-norm preservation and the commutativity of the reals:

$$||Q||_2\,||Q^T||_2 = ||Q^T||_2\,||Q||_2 = ||Q^TQ||_2 = ||I||_2 = 1$$

where the last equality follows assuming we're talking about the induced matrix norm. Thus if P is both orthogonal and a projector, $||P||_2 = 1$

commutivity

However, if we only assume that P is a nonzero projector and lack the supposition that P is orthogonal, observe the following: consider the vector p_1 such that p_1 is the first column vector of matrix P. Since P is nonzero, p_1 is nonzero, or we can choose a column that is. Then since $P^2 = P$, it follows that $Pp_1 = p_1$ by line (3.31) of the week 3 pdf lecture notes. Then it follows that

$$p_1 = Pp_1$$

$$\implies ||p_1||_2 = ||Pp_1||_2$$

$$\implies 1 = \frac{||Pp_1||_2}{||p_1||_2}$$

for a nonzero column vector p_1 of P. Recall that the definition of the induced two norm of a matrix is

$$||P||_2 = \sup_{x\neq 0} \frac{||Px||_2}{||x||_2}$$

Then by the definition of supremum, $||P||_2$ can be no smaller than 1 since we've found a scenario where the norm is in fact 1 for a nonzero vector. That is,

$$||P||_2 \geq 1$$

with equality holding if and only if P is orthogonal in addition, as desired.

□

Please find (b) on the next page...

(b):

Proof. Suppose P is an orthogonal projector. Recall that for real matrices the definition of positive semi-definite requires $x^T Px \geq 0$. This is what we aim to show. To that end, observe that since P is idempotent

$$x^T Px = x^T P^2 x$$

and since P is orthogonal

$$x^T P^2 x = x^T P^T Px = (Px)^T Px$$

According to the lecture notes on matrix norms, the right hand term of the above equality is another way of writing the 2-norm (at least for real valued matrices). That is,

$$(Px)^T (Px) = ||Px||_2 = ||P||_2 \, ||x||_2$$

where again we may split by a property of the two norm. Since in part (a) it was shown that $||P||_2 \geq 1$, and x was taken to be nonzero, it follows that $||P||_2||x||_2 > 0$. Tracing the above equalities all the way back to the start, we have

$$x^T Px > 0$$

meaning P is not only positive semi-definite, but positive definite. Arriving at a stronger claim has me worried - I figure you would have asked for positive definite if it was positive definite.

Next to show the eigenvalues are either zero or one. Consider an eigenvalue-vector pairing λ, v such that

$$Pv = \lambda v$$

Then observe, much like the process in the last part

$$\lambda v^T v = v^T \lambda v = v^T Pv = v^T P^2 v = v^T P^T Pv = (Pv)^T Pv = (\lambda v)^T (\lambda v) - \lambda^2 v^T v$$

Equating the first and last terms we have

$$\lambda v^T v = \lambda^2 v^T v$$

$$\implies \lambda v^T v - \lambda^2 v^T v = 0$$

since $v^T v$ is just some scalar value, we may divide both sides by it giving

$$\lambda - \lambda^2 = 0$$

or

$$\lambda(1 - \lambda) = 0$$

Which implies $\lambda = 0$ or $\lambda = 1$ as desired.

□

Problem 3
Let $A \in \mathbb{R}^{m\times n}$ with $m \geq n$ and let $A = \hat{Q}\hat{R}$ be a reduced QR factorization.

(a) Show that A has rank n if and only if all the diagonal entries of $\hat{R}$ are nonzero.

(b) Suppose $\hat{R}$ has k nonzero diagonal entries for some k with $0 \leq k < n$. What does this imply about the rank of A?

Solution:

(a):

Proof. We can create a series of if and only if statements, proving the claim in one fell swoop. To that end, suppose A has rank n. A has rank $n \iff$ the n columns of A are a linearly independent set of vectors. That is, for $\alpha_i \in \mathbb{R}$ such that

$$\sum_{i=1}^{n} \alpha_i a_i = 0 \tag{3}$$

we require all the α_i to be zero, where a_i denotes the columns of A. This is simply an equivalent definition of linear independence, and as such is an if and only if statement. Since we're working with a reduced QR factorization in both directions, it follows that $\hat{Q}$ has dimension $m \times m$ and $\hat{R}$ has dimension $m \times n$, while also being upper triangular. We can write this decomposition as

$$\hat{Q}\hat{R} = \begin{bmatrix} Q_{11} & \dots & Q_{1m} \\ \vdots & \ddots & \vdots \\ Q_{m1} & \dots & Q_{mn} \end{bmatrix} \begin{bmatrix} R_{11} & \dots & R_{1m} \\ \vdots & \ddots & \vdots \\ 0 & \dots & R_{mn} \end{bmatrix} = \sum_{j-1}^{n} \sum_{i=1}^{m} Q_{ji} r_j$$

For column r_j of matrix $\hat{R}$. However, by equation (3), and since A $= \hat{Q}\hat{R}$, if any of the diagonal entries of the columns r_j are zero, this could allow scalar α_i to be nonzero (as $Q_{ij} = 0$ is possible), violating the definition of linear independence. Therefore we must have that all diagonal entries of $\hat{R}$ must be nonzero. $\square$

(b):

Proof. I claim the rank of A is at most k. We may see this via contradiction. To that end, suppose the rank of $A > k$. We know since $A = \hat{Q}\hat{R}$,

$$Ker(A) = Ker(\hat{Q}\hat{R})$$

which implies

$$\{x : Ax = 0\} = \{x : \hat{Q}\hat{R}x = 0\}$$

by simply applying the definition of kernel. Since $\hat{R}x = 0 \implies \hat{Q}\hat{R}x = 0$ we must have that

$$\ker(\hat{R}) \subseteq \ker(\hat{Q}\hat{R}) = \ker(A)$$

and furthermore, we know that $Dim(ker(\hat{R})) = k$. Invoking the rank nullity theorem, we find

$$Rank(\hat{R}) + Nullity(\hat{R}) = dim(\mathbb{R}^n)$$

and so

$$Rank(\hat{R}) + k = n$$

Then, $Rank(A) = n$ and

$$Rank(A) = k - Rank(\hat{R})$$

but we supposed that $Rank(A) > k$, which is impossible with the above statement. We have a contradiction, and as such it must be that instead $Rank(A) \leq k$, as desired. $\square$

Problem 4
Determine the (i) eigenvalues, (ii) determinant, and (ii) singular values of a Householder reflector. Give a geometric and algebraic argument for the the eigenvalues.

Solution: **(i):** First let's examine the eigenvalues. Algebraically, recall that eigenvalues arise when the application of a linear transformation to a vector results in a constant scalar times that vector. Geometrically, this means the linear transformation only scales (by a constant) or reorients our vector (with a negative value). For instance, a linear transformation that doubles the standard basis vector and does not change its direction at all would have an eigenvalue of 2.

From the lecture notes, a Householder reflector is a linear transformation that is symmetric and orthogonal. The name reflector gives it away - as a result of orthogonality the absolute values of our eigenvalues will be 1. Since this linear transformation reflects our original vector without any scaling, it either leaves the vector alone (eigenvalue of 1) or it reflects the vector within a hyperplane (eigenvalue of -1). The eigenvectors that are invariant will be perpendicular to the hyperplane, while those exactly reflected are found in the plane.

Now let's see this algebraically. Take an eigenvector value pair of the householder reflector H such that v is a unit normal vector. Then computing eigenvalues using the definition of H

$$det(H - \lambda I) = det(I - 2vv^T - \lambda I) = 0$$

$$\implies (1 - 2v_{11}^2)(1 - 2v_{22}^2)\dots(1 - 2v_{nn}^2) = 0$$

Thus we see the eigenvalues are given by $\lambda_i = 1 - 2v_{ii}^2$, and we get one for each n. Since v was chose to be a unit normal vector, we have that $v_{ii}^2 = 1 \ \forall n$.

$$\therefore \lambda_i = \pm 1$$

as desired.

(ii): In general it is known that the determinant of a matrix is equal to the product of its eigenvalues. From the above argument, all eigenvalues are either 1 or -1, and so the determinant of a Householder reflector could only ever be ± 1. However, any reflection that preserves length (if I'm not mistaken) should have a determinant of -1.

(iii): Let's keep the geometric ball rolling. Singular values, at least if H acts on $\mathbb{R}^n$, can be visualized as the length of the principal axes after application of H. That is, we take the unit hyper ball, hit it with H, then see what the lengths of the semi-axes are - these are our singular values. Hey! H is just a reflector and won't change the lengths of these principal axes - in other words it won't distort the unit hyper ball. As such, it must be that all the real singular values are 1.

Problem 5
Let $A \in \mathbb{R}^{m\times n}$. Show that $\text{Cond}(A^T A) = (\text{Cond}(A))^2$.

Solution: Note that in general, the relative condition number k is bounded above by $||A||\,||A^{-1}||$. However, in the case of the 2-norm, the relative condition number is exactly $||A||\,||A^{-1}||$. That is,

$$k(x_0) = ||A||_2\,||A^{-1}||_2 = \text{Cond(A)}$$

for some point x_0. We also know that in general

$$||A||_2 = \sqrt{\max_\lambda(A^*A)} = \sqrt{\max_\lambda(A^T A)}$$

for real valued A, per the lecture notes on matrix norms.

Then

$$\begin{aligned}
\text{Cond}(A^T A) &= ||A^T A||_2 \;\; ||(A^T A)^{-1}||_2 \\
\text{Cond}(A^T A) &= ||\max_\lambda(A^T A)||_2 \;\; ||\max_\lambda(A^T A)^{-1}||_2 \\
&= ||\left(\sqrt{\max_\lambda(A^T A)}\right)^2||_2 \;\; ||\left(\sqrt{\max_\lambda(A^T A)^{-1}}\right)^2||_2 \qquad (4) \\
&= ||A||_2^2 \;\; ||A^{-1}||_2^2 \\
&= \text{Cond}(A)^2
\end{aligned}$$

Where the second line follows since the relative condition number is bounded above by, or in this case is exactly, $||A||_2||A^{-1}||_2$, coupled with the equivalent max λ definition.