1. Let $\mathbf{A}$ and $\mathbf{B}$ both be $n \times n$ and assume that there is an invertible matrix $\mathbf{S}$ such that $\mathbf{SAS}^{-1}$ and $\mathbf{SBS}^{-1}$ are both diagonal (but not necessarily equal). Prove that $\mathbf{AB} = \mathbf{BA}$.

   Solution: Let

   $$\mathbf{SAS}^{-1} = \mathbf{D}_A \text{ and } \mathbf{SBS}^{-1} = \mathbf{D}_B.$$

   Write

   $$\mathbf{A} = \mathbf{S}^{-1}\mathbf{D}_A\mathbf{S} \text{ and } \mathbf{B} = \mathbf{S}^{-1}\mathbf{D}_B\mathbf{S}.$$

   Note

   $$\mathbf{AB} = \mathbf{S}^{-1}\mathbf{D}_A\mathbf{S}\mathbf{S}^{-1}\mathbf{D}_B\mathbf{S} = \mathbf{S}^{-1}\mathbf{D}_A\mathbf{D}_B\mathbf{S} = \mathbf{S}^{-1}\mathbf{D}_B\mathbf{D}_A\mathbf{S} = \mathbf{S}^{-1}\mathbf{D}_B\mathbf{S}\mathbf{S}^{-1}\mathbf{D}_A\mathbf{S} = \mathbf{BA}.$$

2. Suppose that $\mathbf{A}$ is symmetric and positive definite, while $\mathbf{B}$ is symmetric. Prove that there is a real invertible matrix $\mathbf{S}$ such that

   $$\mathbf{A} + \mathbf{B} = \mathbf{S}\left(\mathbf{I} + \mathbf{D}\right)\mathbf{S}^T$$

   where $\mathbf{D}$ is diagonal.

   Solution: Since $\mathbf{A}$ is SPD it has a Cholesky factorization $\mathbf{A} = \mathbf{LL}^T$. Factor out the Cholesky factors from the sum:

   $$\mathbf{A} + \mathbf{B} = \mathbf{L}\left(\mathbf{I} + \mathbf{L}^{-1}\mathbf{BL}^{-T}\right)\mathbf{L}^T.$$

   Notice that the matrix $\mathbf{L}^{-1}\mathbf{BL}^{-T}$ is real and symmetric, so it has a real orthogonal eigenvalue decomposition

   $$\mathbf{L}^{-1}\mathbf{BL}^{-T} = \mathbf{QDQ}^T.$$

   Insert this and simplify

   $$\mathbf{L}\left(\mathbf{I} + \mathbf{QDQ}^T\right)\mathbf{L}^T = \mathbf{L}\left(\mathbf{QQ}^T + \mathbf{QDQ}^T\right)\mathbf{L}^T = \mathbf{LQ}\left(\mathbf{I} + \mathbf{D}\right)\left(\mathbf{LQ}\right)^T.$$

   The matrix $\mathbf{S} = \mathbf{LQ}$ is real and invertible.

3. Suppose that $\mathbf{A}$ and $\mathbf{B}$ are both $n \times n$. Prove that $\mathbf{AB}$ and $\mathbf{BA}$ have the same eigenvalues.

   Solution: Suppose that $\mathbf{AB}\vec{v} = \lambda\vec{v}$ with $\vec{v} \neq \vec{0}$. Multiply by $\mathbf{B}$: $\mathbf{BAB}\vec{v} = \lambda\mathbf{B}\vec{v}$. Define $\vec{w} = \mathbf{B}\vec{v}$. Then

   $$\mathbf{BA}\vec{w} = \lambda\vec{w}.$$

   This equation only implies that λ is an eigenvalue of $\mathbf{BA}$ when $\vec{w} \neq \vec{0}$. So what we've shown so far is

   > If $\mathbf{AB}$ has eigenvalue λ and eigenvector $\vec{v}$ then $\mathbf{BA}$ also has eigenvector λ provided that $\mathbf{B}\vec{v} \neq \vec{0}$.

   So what happens when $\vec{v} \neq \vec{0}$ but $\mathbf{B}\vec{v} = \vec{0}$? In that case $\lambda = 0$ must be an eigenvalue of $\mathbf{AB}$ because $\mathbf{AB}\vec{v} = \mathbf{A}\vec{0} = \vec{0} = (0)\vec{v}$. So we can conclude at this point that

   > If $\mathbf{AB}$ has eigenvalue $\lambda \neq 0$ and eigenvector $\vec{v}$ then $\mathbf{BA}$ also has eigenvector λ.

By re-labeling $\mathbf{A}$ and $\mathbf{B}$ we find that

The nonzero eigenvalues of $\mathbf{AB}$ and $\mathbf{BA}$ are the same.

To complete the proof we need to show that if $\lambda = 0$ is an eigenvalue of $\mathbf{AB}$ then it must also be an eigenvalue of $\mathbf{BA}$. Clearly if $\mathbf{A}$ and $\mathbf{B}$ are both invertible then it's not possible for $\lambda = 0$ to be an eigenvalue of $\mathbf{AB}$ or $\mathbf{BA}$. If at least one of the factors is singular then the product is also singular, meaning that it has a zero eigenvalue regardless of the order of the factors. This completes the proof.

4. Suppose that $\mathbf{A}$and $\mathbf{B}$ are both symmetric and positive definite. Prove that $\mathbf{AB}$ has real, positive eigenvalues. (Hint: Show that $\mathbf{AB}$ is similar to a symmetric and positive definite matrix.)

Solution: Since $\mathbf{A}$ is SPD it has a Cholesky factorization

$$\mathbf{A} = \mathbf{L}\mathbf{L}^T$$

so

$$\mathbf{AB} = \mathbf{L}\mathbf{L}^T\mathbf{B}.$$

The Cholesky factor $\mathbf{L}$ is invertible. Multiply from the left by $\mathbf{L}^{-1}$ and from the right by $\mathbf{L}$ to get

$$\mathbf{L}^{-1}\mathbf{ABL} = \mathbf{L}^T\mathbf{BL}.$$

The equation above says that $\mathbf{AB}$ is similar to $\mathbf{L}^T\mathbf{BL}$, so they have the same eigenvalues. The matrix $\mathbf{L}^T\mathbf{BL}$ is symmetric, so it has real eigenvalues. Also note that

$$\vec{x}^T\mathbf{L}^T\mathbf{BL}\vec{x} = \vec{y}^T\mathbf{B}\vec{y} \geq 0 \; \forall \; \vec{y} \neq \vec{0}.$$

Since $\vec{y} = \vec{0}$ only when $\vec{x} = \vec{0}$ (because $\mathbf{L}$ is invertible), we have shown that $\mathbf{L}^T\mathbf{BL}$ is positive definite, and therefore has real eigenvalues.

5. Suppose that $\mathbf{L} \in \mathbb{R}^{n\times n}$ is a symmetric matrix with eigenvalues $\lambda_1, \ldots, \lambda_m$. (Note that m can be $\leq n$.) Prove that

$$\mathbf{L} = \lambda_1\mathbf{P}_1 + \cdots + \lambda_m\mathbf{P}_m$$

where $\mathbf{P}_i$ is an orthogonal projection matrix that projects orthogonally onto the nullspace of $\mathbf{L} - \lambda_i\mathbf{I}$.

Solution: Since $\mathbf{A}$ is symmetric it has a real orthogonal eigenvalue decomposition

$$\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T.$$

Write

$$\mathbf{Q} = [\vec{q}_1, \cdots, \vec{q}_n]. \ \ \mathbf{\Lambda} = \text{diag}(\mu_1, \ldots, \mu_n).$$

The eigenvalue decomposition can be equivalently written

$$\mathbf{A} = \mu_1 \vec{q}_1 \vec{q}_q^T + \ldots + \mu_n \vec{q}_n \vec{q}_n^T.$$

If we group the eigenvalues we get

$$\mathbf{A} = \lambda_1 [\vec{q}_1, \cdots][\vec{q}_1, \cdots]^T + \ldots + \lambda_m [\vec{q}_r, \cdots][\vec{q}_r, \cdots]^T.$$

Each of the matrices $[\vec{q}_r, \cdots][\vec{q}_r, \cdots]^T$ is an orthogonal projection matrix that projects orthogonally onto the eigenspace associated with λ_r. The notation in this solution is not great, but making it precisely correct requires an overload of notation.

6. Prove that if $\mathbf{A}$ is invertible (and, for simplicity, real), then there is an orthogonal matrix $\mathbf{U}$ and a symmetric and positive definite matrix $\mathbf{E}$ such that $\mathbf{A} = \mathbf{UE}$. This is analogous to writing complex numbers in polar form $x + \mathrm{i}y = \rho e^{\mathrm{i}\theta}$. Show that if $\mathbf{A}$ is square but not invertible then there is a similar factorization but with $\mathbf{E}$ being symmetric and non-negative definite.

   Solution: Consider the SVD of $\mathbf{A}$:

   $$\mathbf{A} = \mathbf{P\Sigma Q}^T = \mathbf{PQ}^T\mathbf{Q\Sigma Q}^T = \mathbf{UE}$$

   where

   $$\mathbf{U} = \mathbf{PQ}^T \text{ and } \mathbf{E} = \mathbf{Q\Sigma Q}^T.$$

   The diagonal entries of $\mathbf{\Sigma}$ are the eigenvalues of $\mathbf{E}$. When $\mathbf{A}$ is invertible Σ has positive diagonal entries, and when $\mathbf{A}$ is not invertible $\mathbf{\Sigma}$ has some positive and some zero diagonal entries. The matrix $\mathbf{U}$ is a product of orthogonal matrices and is therefore orthogonal.

7. Let $\mathbf{A} = \mathbf{S\Lambda S}^{-1}$ be diagonalizable. Find a symmetric positive definite matrix $\mathbf{K}$ that defines an inner product $\langle \vec{x}, \vec{y} \rangle = \vec{x}^T\mathbf{K}\vec{y}$ such that $\mathbf{A}$ commutes with its $\mathbf{K}$-adjoint.

   Solution: First recall that the $\mathbf{K}$-adjoint of $\mathbf{A}$ is

   $$\mathbf{A}^\dagger = \mathbf{K}^{-1}\mathbf{A}^T\mathbf{K}.$$

   We will draw inspiration from the idea that 'commutes with adjoint' is a generalization of the definition of a normal matrix, and that the eigenvectors of normal matrices are orthogonal. So we want to find an SPD matrix $\mathbf{K}$ such that the Gram matrix formed using the eigenvectors of $\mathbf{A}$ and the inner product defined by $\mathbf{K}$ is the identity:

   $$\mathbf{S}^T\mathbf{KS} = \mathbf{I}.$$

   Consider the SVD of $\mathbf{S}$:

   $$\mathbf{S} = \mathbf{P\Sigma Q}^T$$

where all factors are square and invertible. If we set

$$\mathbf{K} = \mathbf{P}\mathbf{\Sigma}^{-2}\mathbf{P}^T$$

then we find

$$\mathbf{S}^T\mathbf{K}\mathbf{S} = \mathbf{Q}\mathbf{\Sigma}\mathbf{P}^T\mathbf{P}\mathbf{\Sigma}^{-2}\mathbf{P}^T\mathbf{P}\mathbf{\Sigma}\mathbf{Q}^T = \mathbf{I},$$

as desired.

We should check that our intuition has guided us correctly though, i.e. does $\mathbf{A}$ commute with its $\mathbf{K}$-adjoint? First notice that the above equation implies

$$\mathbf{S}^{-1}\mathbf{K}^{-1}\mathbf{S}^{-T} = \mathbf{I}$$

which implies

$$\mathbf{K}^{-1}\mathbf{S}^{-T} = \mathbf{S}.$$

We also have

$$\mathbf{S}^T\mathbf{K} = \mathbf{S}^{-1}.$$

Now simplify

$$\mathbf{A}^\dagger\mathbf{A} = \mathbf{K}^{-1}\mathbf{A}^T\mathbf{K}\mathbf{A} = \mathbf{K}^{-1}\mathbf{S}^{-T}\mathbf{\Lambda}\mathbf{S}^T\mathbf{K}\mathbf{S}\mathbf{\Lambda}\mathbf{S}^{-1} = \mathbf{K}^{-1}\mathbf{S}^{-T}\mathbf{\Lambda}^2\mathbf{S}^{-1} = \mathbf{S}\mathbf{\Lambda}^2\mathbf{S}^{-1} = \mathbf{A}^2$$

and

$$\mathbf{A}\mathbf{A}^\dagger = \mathbf{A}\mathbf{K}^{-1}\mathbf{A}^T\mathbf{K} = \mathbf{S}\mathbf{\Lambda}\mathbf{S}^{-1}\mathbf{K}^{-1}\mathbf{S}^{-T}\mathbf{\Lambda}\mathbf{S}^T\mathbf{K} = \mathbf{S}\mathbf{\Lambda}^2\mathbf{S}^T\mathbf{K} = \mathbf{S}\mathbf{\Lambda}^2\mathbf{S}^{-1} = \mathbf{A}^2.$$

We have therefore found that every diagonalizable matrix is normal with respect to some inner product.

8. Let $\mathbf{A} \in \mathbb{R}^{m\times n}$ have full column rank. Prove that $\|\mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\|_2 = 1/\sigma_n$ where σ_n is the n^{th} singular value of $\mathbf{A}$. (Note that $\mathbf{A}^{-1}$ is not defined, so your proof cannot use $\mathbf{A}^{-1}$.)

   Solution: Consider the SVD of $\mathbf{A}$:

   $$\mathbf{A} = \mathbf{P}\mathbf{\Sigma}\mathbf{Q}^T$$

   where $\mathbf{Q}$ is orthogonal, $\mathbf{\Sigma}$ is diagonal and invertible, and $\mathbf{P}$ is non-square. Plug in and simplify:

   $$\mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1} = \mathbf{P}\mathbf{\Sigma}\mathbf{Q}^T(\mathbf{Q}\mathbf{\Sigma}\mathbf{P}^T\mathbf{P}\mathbf{\Sigma}\mathbf{Q}^T)^{-1} = \mathbf{P}\mathbf{\Sigma}\mathbf{Q}^T(\mathbf{Q}\mathbf{\Sigma}^2\mathbf{Q}^T)^{-1} = \mathbf{P}\mathbf{\Sigma}\mathbf{Q}^T\mathbf{Q}\mathbf{\Sigma}^{-2}\mathbf{Q}^T = \mathbf{P}\mathbf{\Sigma}^{-1}\mathbf{Q}^T.$$

   This is an SVD[1] of $\mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}$. The singular values of $\mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}$ are one over the singular values of $\mathbf{A}$. The 2-norm of $\mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}$ is the largest singular value of $\mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}$, which is one over the smallest singular value of $\mathbf{A}$, which is $1/\sigma_n$.

[1] Technically we need to insert permutation matrices to re-order the singular values so that they decrease along the diagonal.