1. Let A be a real matrix with distinct eigenvalues $\lambda_1, \ldots, \lambda_n$, which are real and positive. How many real matrices B satisfy the matrix equation $A = B^2$.

   Answer: Any square matrix is similar to an upper triangular matrix J (for example its Jordan form). Then

   $$A = B^2 = PJP^{-1}PJP^{-1} = PJ^2P^{-1},$$

   so A is similar to J^2.

   By question 4 of homework 4, we know that $\det(\lambda I - J) = \prod_{k=1}^{n}(\lambda - J_{k,k})$, so the eigenvalues of B are $\mu_k = J_{k,k}$. The eigenvalues of A are $\lambda_k = J_{k,k}^2$, because the square of an upper triangular matrix is an upper triangular matrix with the diagonal entries squared. Thus each eigenvalue of B must be the square root of a distinct eigenvalue of A.

   We thus have that B has distinct eigenvalues, and hence a basis of genuine eigenvectors $\{v_1, \ldots, v_n\}$. Also, since

   $$Av_k = B(Bv_k) = B(\mu_k v_k) = \mu_k^2 v_k = \lambda_k v_k,$$

   these eigenvectors must also be eigenvectors of A. Thus the only freedom in the choice of B is the freedom to choose the sign of the eigenvalue for each v_k. Since there are n v_k's we see that there are 2^n possible B's.

   Note: An alternative proof of this fact is to note that since A has distinct eigenvalues, and so $\mathbb{R}^n$ may be decomposed into a sum of one-dimensional eigenspaces of A. You can then prove that B maps the eigenspaces into themselves, and use this to determine B up to signs, and then count the number of signs. This proof has the advantage of giving more intuition about the behaviour of these matrices, but I find proof presented above to be cleaner.

2. A matrix A is said to be monotone if for $x = (x_1, \ldots, x_n)^T$, $Ax \geq 0$ (i.e. each element of Ax is ≥ 0) implies $x \geq 0$.

   (a) Show that A is invertible.

   Answer: Suppose $Ax = 0 \geq 0$, then $x \geq 0$. By linearity, we also have that $A(-x) = 0 \geq 0$, so $-x \geq 0$. Thus $x = 0$. We thus have that A has a only a trivial nullspace. Since A is square, this implies that it is invertible.

   (b) Show that A is monotone if and only if A^{-1} exists and all its entries are non-negative.

   Answer: First suppose A is monotone. Then A^{-1} exists by part a). Also note that if $\tilde{a}_k$ is the kth column of A^{-1}, then

   $$AA^{-1} = I \Rightarrow A\tilde{a}_k = e_k \geq 0,$$

   where e_k is the kth standard basis vector of $\mathbb{R}^n$, which has non-negative entries. The monotonicity of A then gives that $\tilde{a}_k \geq 0$. Thus each column of A^{-1} has non-negative entries, and so we are done.

   We now suppose that A^{-1} exists and all its entries are non-negative. Since A^{-1} is invertible, we have that the $\tilde{a}_k$'s form a basis.

   $$Ax = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \Rightarrow x = A^{-1}\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \Rightarrow x = \sum_{k=1}^{n} b_k \tilde{a}_k.$$

Since each element of $\tilde{a}_k$ is non-negative, we have that $b_k \geq 0$ for all k implies that $x \geq 0$. This is the definition of A being monotone.

3. Let A and B be rectangular matrices of size $n \times m$ and $m \times n$ respectively. Prove that the non-zero eigenvalues of AB and BA are the same. Express the corresponding eigenvectors of AB in terms of those of BA. Show that if $\lambda + AB$ is invertible for some $\lambda \neq 0$, then

$$\frac{\lambda}{\lambda + AB} + A\frac{1}{\lambda + BA}B = I.$$

Answer: Let v be an eigenvector of BA with eigenvalue $\lambda \neq 0$, then

$$(AB)Av = A(BAv) = \lambda Av.$$

Since $\lambda \neq 0$, this tells us that Av is an eigenvector of AB with eigenvalue of λ. Thus all non-zero eigenvalues of BA are also eigenvalues of AB.

By switching the roles of A and B above, we may also see that non-zero eigenvalues of AB are also eigenvalues of BA.

Now suppose that $\lambda + AB$ is invertible for $\lambda \neq 0$ (and so $\lambda + BA$ is invertible, by the above) Note that $(\lambda + AB)A = \lambda A + ABA = A(\lambda + BA)$, so $(\lambda + AB)A(\lambda + BA)^{-1} = A$. Thus

$$\begin{aligned}\lambda + (\lambda + AB)A(\lambda + BA)^{-1}B &= \lambda + AB \\ \Rightarrow (\lambda + AB)^{-1}\lambda + (\lambda + AB)^{-1}(\lambda + AB)A(\lambda + BA)^{-1}B &= I \\ \Rightarrow \frac{\lambda}{\lambda + AB} + A\frac{1}{\lambda + BA}B &= I.\end{aligned}$$

4. Let C be an $n \times n$ matrix.

   (a) Show that $\text{tr}C = 0$ if and only if $C = SDS^{-1}$ for some invertible S and some matrix D with all diagonal entries equal to 0.

   Answer: First note that if $C = SDS^{-1}$, for such a D, then

$$\text{tr}C = \text{tr}[SDS^{-1}] = \text{tr}[S^{-1}SD] = \text{tr}D = \sum_{i=1}^{n} D_{i,i} = 0.$$

   We now use induction to prove the other direction. Clearly if $n = 1$ and $\text{tr}C = 0$, then $C = 0$, so the result holds. Now suppose that the result holds for any $k < n$.

   We claim that C is similar to a matrix of the form $\begin{pmatrix} 0 & b^T \\ a & C_1 \end{pmatrix}$. If we accept this claim, then the induction hypothesis says that C_1 is similar to a matrix D_1 with zeros on the diagonal, since $\text{tr}C_1 = \text{tr}C = 0$. We may then the last $n-1$ vectors in the basis that C is written in to write

$$P^{-1}CP = \begin{pmatrix} 0 & \tilde{b}^T \\ \tilde{a} & D_1 \end{pmatrix},$$

   which is a matrix with only zeros on the diagonal.

   Proof of claim: Since $\text{tr}C = 0$, we either have that $C = 0$ (and so we are done) or $C \neq \lambda I$ for any λ (since then $\text{tr}C = n\lambda \neq 0$). If $C \neq 0$, then there exists a $v_1 \in \mathbb{R}^n$ such that $v_2 = Cv_1 \neq \lambda v_1$ for any λ. We may then choose $v_3, \dots, v_n$ to complete a basis for $\mathbb{R}^n$. If we write C in this basis then C will have a 0 in the top left corner because Cv_1 has no component in the v_1 direction.

(b) Show that tr$C = 0$ if and only if $C = AB - BA$ for some $n \times n$ matrices A and B.

Answer: If $C = AB - BA$, the tr$C = \mathrm{tr}AB - \mathrm{tr}BA = 0$.
To prove the other direction, let A be some diagonal matrix with the $A_{j,j}$'s all distinct. Then asking $C = AB - BA$ is the same as asking that

$$C_{i,j} = A_{i,i}B_{i,j} - B_{i,j}A_{j,j},\ \forall i,j$$
$$\Rightarrow B_{i,j} = \frac{C_{i,j}}{A_{i,i} - A_{j,j}},\ \forall i,j$$

which is well defined, so we have found our A and B.
Note: It is also possible to prove this induction. This uses the claim from part a) and the fact that if $\begin{pmatrix} 0 & b^T \\ a & C_1 \end{pmatrix} = \tilde{A}\tilde{B} - \tilde{B}\tilde{A}$, then

$$C = P\tilde{A}P^{-1}P\tilde{B}P^{-1} - P\tilde{B}P^{-1}P\tilde{A}P^{-1} = AB - BA,$$

so C is also commutator of two square matrices.

5. Show that if A is a $k \times k$ matrix with complex entries and $A^n = I$ for some $n > 0$, then A has a basis of eigenvectors.

Answer: There are two ways to prove this, by proving that A has no eigenvectors that are generalized and not genuine, or by using the minimal polynomial of A.

Eigenvector approach: Suppose that A has at least one eigenvector that is generalized and not genuine, then A has a Jordan chain of size at least 2, i.e. there exists a v such that $v_1 = (A - \lambda I)v \neq 0$, but $(A - \lambda I)^2 v = 0$. Note that v and v_1 are linearly independent. Then,

$$Av = \lambda v + v_1$$
$$\Rightarrow A^2 v = \lambda(\lambda v + v_1) + \lambda v_1$$
$$\vdots$$
$$A^n v = \lambda^n v + n\lambda^{n-1} v_1. \tag{1}$$

Since $A^n = I$, we have that $(1 - \lambda^n)v = n\lambda^{n-1}v_1$. Since v and v_1 are linearly independent, both sides are zero. Thus $\lambda^n = 1$ and $\lambda^{n-1} = 0$, which is a contradiction.

Note: It is not true that any eigenvector that is generalized and not genuine satisfies (1). As an example take the matrix $A = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$. Then $Ae_3 = \lambda e_3 + e_2$, and $A^2 e_3 = \lambda^2 e_3 + 2\lambda e_2 + e_1$, which is not the same behaviour as we used in the proof.

Minimal polynomial approach: If we define $p(x) = x^n - 1$, then we have that $p(A) = 0$. **Note that this does not imply that p is the characteristic polynomial of A, even if it had the right degree.** , $p(x) = x^n - 1 = \prod_{k=1}^{n}(x - \omega_k)$, where ω_k are the nth roots of 1. We see that each term only appears once, so all roots of p have multiplicity 1.

It is a fact that $p(A) = 0$ implies that the minimal polynomial of A must divide p. Since the roots of the minimal polynomial are eigenvalues of A, we have that all of the eigenvalues of A must be nth roots of 1. More importantly, we have that the multiplicity of every root of the minimal polynomial has multiplicity 1, which tells us that the Jordan normal form of A has only Jordan blocks of size 1. Thus A is diagonalizable and has a basis of eigenvectors.

6. A square matrix S is stochastic if all its elements are non-negative and the sum of the elements in each column is 1.

$$\text{i.e. } \sum_{i=1}^{n} S_{i,j} = 1, 1 \leq j \leq n.$$

Show that

(a) $\lambda = 1$ is an eigenvalue of S.

Answer: Note that

$$S^T \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} S_{i,1} \\ \vdots \\ \sum_{i=1}^{n} S_{i,n} \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix},$$

so 1 is an eigenvalue of S^T. Since $\det(S - I) = \det(S^T - I) = 0$, we also have that 1 is an eigenvalue of S.

(b) All eigenvalues λ_i of S lie in the closed unit disk. i.e. $|\lambda_i| \leq 1$ for all i.

Answer: We again note that all of the eigenvalues of S are also eigenvalues of S^T. Now suppose that v is an eigenvector of S^T. Then

$$|\lambda||v_i| = |\lambda v_i| = \left| \sum_{i=1}^{n} S_{i,j} v_j \right| \leq \left| \sum_{i=1}^{n} S_{i,j} \right| \max_{1 \leq j \leq n} |v_j| = \max_{1 \leq j \leq n} |v_j|.$$

Choosing i such that $|v_i| = \max_{1 \leq j \leq n} |v_j|$ gives that $|\lambda| \leq 1$.