(1) (a) An $n \times n$ matrix M is Toeplitz if and only if these equations are satisfied:

$$M_{i,j} = M_{i+k,j+k}$$

for all positive integers i, j, k such that $i+k, j+k \leq n$. Since addition and scalar multiplication of matrices occurs elementwise, any linear combination of Toeplitz matrices will still be Toeplitz, because these equations are linear. Therefore, $n \times n$ Toeplitz matrices are a vector space.

(b) An $n \times n$ Toeplitz matrix is specified by $2n-1$ independent real numbers, so the dimension of this space is $2n-1$.

(c) $M \mapsto M^\top$ works.

(d) No. A counterexample is given by

$$M = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

This problem is tricky because there is no 1×1 or 2×2 counterexample.

(e) Yes. They are the intersection of the space of Toeplitz matrices with the space of symmetric matrices. The intersection of two vector subspaces is a vector subspace.

(2) The answer is $-(A^{-2})^\top$.

To see this, treat the entries of A as variables, and consider the matrix of differentials

$$dA := \begin{pmatrix} da_{11} & da_{12} & \cdots \\ da_{21} & da_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

as usual. Let's apply $d(-)$ to the matrix identity $AA^{-1} = \mathrm{Id}$. On the one hand,

$$d(\mathrm{Id}) = 0$$

since Id doesn't change when the entries of A change. On the other hand, the Leibniz rule implies

$$d(AA^{-1}) = (dA)A^{-1} + A(dA^{-1}).$$

Thus, we conclude that

$$0 = (dA)A^{-1} + A(dA^{-1}),$$

which implies that $dA^{-1} = -A^{-1}(dA)A^{-1}$.

Next, we apply this to $f(A)$. By linearity and cyclic invariance of trace, we have

$$\begin{aligned} d(f(A)) &= d(\mathrm{trace}(A^{-1})) \\ &= \mathrm{trace}(dA^{-1}) \\ &= \mathrm{trace}(-A^{-1}(dA)A^{-1}) \\ &= \mathrm{trace}(-(dA)A^{-2}). \end{aligned}$$

The coefficient of da_{ij} in this trace is $-(A^{-2})_{ji}$. (This is the (j,i)-th entry of the matrix $-A^{-2}$.) Therefore, the gradient of $f(A)$ is $-(A^{-2})^\top$.

Concretely, this means that if the (i,j)-th entry of A is modified by adding small real number ϵ, while all other entries of A remain fixed, then the value $f(A)$ is modified by adding the real number $-(A^{-2})_{ji}\,\epsilon$.

(3) (a) False. We have

$$\exp\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

which is not symmetric. We remark that e^A is not even defined when A is not square.

(b) True. If A is a (real) symmetric matrix, we can write $A = VDV^\top$ where V is orthogonal and D is diagonal. Then
$$e^A = Ve^DV^\top.$$
The matrix e^D is positive definite because it is diagonal with all diagonal entries *positive*, and conjugating by the orthogonal matrix V preserves this property.

(c) False. We have
$$\exp(-1) = (e^{-1}),$$
and (-1) is orthogonal but (e^{-1}) is not.

(d) The convention, used on Wikipedia and in Strang's book, is that a 'positive definite matrix' is implicitly symmetric. Therefore, the present statement is **true** because statement (b) is true.

Remark. If one drops this convention, then the answer becomes **false**. A student deserves at least partial credit if they find a counterexample which demonstrates this. One example is
$$A = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix},$$
which does satisfy $x^\top Ax \geq 0$ (for $x \in \mathbb{R}^2$) with equality if and only if $x = 0$. On the other hand, writing $A = \mathrm{Id} + 2\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ allows one to compute that
$$e^A = e\begin{pmatrix} \cos(2) & -\sin(2) \\ \sin(2) & \cos(2) \end{pmatrix}.$$
This matrix does not satisfy $x^\top e^A x \geq 0$, as can be seen by taking $x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

(e) True. If D is diagonal with entries $d_1, d_2, \ldots$, then e^D is diagonal with entries $e^{d_1}, e^{d_2}, \ldots$.

(4) This is proved in Section 6.4 of Strang's book. We rewrite the proof here.

Let S be a symmetric matrix, and let x be an eigenvector with eigenvalue λ. *A priori*, x may have complex entries and λ may be a complex number. Let $\overline{x}$ be the vector whose entries are the complex conjugates of the entries of x.

Let's compute $\overline{x}^\top Sx$ in two different ways.

First, $Sx = \lambda x$ implies that $\overline{x}^\top Sx = \lambda\, \overline{x}^\top x$.

Second, taking conjugate-transpose of $Sx = \lambda x$ yields $\overline{x}^\top S = \overline{\lambda}\overline{x}^\top$. (We have used that S is real and symmetric, so $\overline{S}^\top = S$.) Right-multiplying by x yields $\overline{x}^\top S = \overline{\lambda}\, \overline{x}^\top x$.

Thus, we conclude that $\lambda\, \overline{x}^\top x = \overline{\lambda}\, \overline{x}^\top x$.

We would like to divide by $\overline{x}^\top x$. We may do so because this is a positive real number: it is the sum of squared magnitudes of the complex number entries of x, which is nonzero because x is nonzero (since it's an eigenvector). We conclude that $\lambda = \overline{\lambda}$, so λ is real.

(5) Since S is real and symmetric, we can write $S = VDV^\top$ where V is orthogonal and D is diagonal. The problem allows us to assume that each diagonal entry of D is positive.

Lemma. For $y \in \mathbb{R}^n$, we have $y^\top Dy \geq 0$ with equality if and only if $y = 0$.

Proof. If $y_1, \ldots, y_n$ are the entries of y, and $d_1, \ldots, d_n$ are the diagonal entries of D, then
$$y^\top Dy = d_1y_1^2 + d_2y_2^2 + \cdots + d_ny_n^2.$$
The claim follows from the hypothesis that all the d_i are positive. □

We deduce that
$$x^\top VDV^\top x = (V^\top x)^\top D(V^\top x) \geq 0,$$

with equality if and only if $V^\top x = 0$, i.e. if $x = 0$. 6

(6) (a) False. The matrix $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ has a singular value of 1, but all eigenvalues are zero.

(b) False. The matrix from (a) has one nonzero singular value, but no nonzero eigenvalue.

(c) True. The best way to prove this is to observe that, for any matrix A, there are diagonalizable matrices M which are arbitrarily close to A. (See the discussion on page 343 of Section 6.4 of Strang's book.) If M is diagonalizable, we can write $M = VDV^{-1}$ where the diagonal entries of D are the eigenvalues of M. Then the eigenvalues of $M^2 = VD^2V^{-1}$ are the squares of the eigenvalues of M. As M approaches A, this conclusion transfers to A because eigenvalues vary continuously with the entries of the matrix.

(d) False. The matrix from (a) has singular values $\{0, 1\}$, but its square is $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ which has singular values $\{0, 0\}$.

(e) True. If A is invertible, its SVD looks like $A = U\Sigma V^\top$ where U and V are square orthogonal matrices and Σ is an invertible diagonal matrix. Then $A^{-1} = V\Sigma^{-1}U^\top$ is the SVD for A^{-1}. The claim follows from the fact that Σ^{-1} is the diagonal matrix whose diagonal entries are the reciprocals of those of Σ.

(f) False. The matrix

$$\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{2} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

is Markov, but the SVD shown above indicates that its singular values are $\{\sqrt{2}, 0\}$. If (following Strang) one wants a Markov matrix to also have strictly positive entries, then one obtains a counterexample by perturbing the above matrix by a small amount:

$$\begin{pmatrix} 1-\epsilon & 1-\epsilon \\ \epsilon & \epsilon \end{pmatrix}$$

Since the singular values vary continuously with the entries of the matrix, for sufficiently small $\epsilon > 0$, the singular values of this matrix will also not be equal to 1.

This problem is tricky because any *symmetric* Markov matrix does have a singular value equal to 1, because, if a matrix is symmetric, then its singular values and eigenvalues coincide.

(g) True.

Lemma. Any eigenvalue of a square orthogonal matrix satisfies $|\lambda| = 1$.

Proof. [1] If U is orthogonal, then $U^\top U = \text{Id}$ by definition. Let x be an eigenvector of U, with eigenvalue λ. As in the solution to (4), taking conjugate-transpose of $Ux = \lambda x$ gives $\overline{x}^\top U^\top = \overline{\lambda}\overline{x}^\top$. (We used that U has real entries.) Multiplying these equations together implies

$$\overline{x}^\top U^\top U x = |\lambda|^2 \overline{x}^\top x.$$

The LHS equals $\overline{x}^\top x$. As in (4), $\overline{x}^\top x$ is a positive real number, so we can divide by it to find that $1 = |\lambda|^2$. □

The eigenvalues of a 3×3 orthogonal matrix are the roots of a cubic polynomial with real coefficients. These roots appear in complex conjugate pairs, so at least one root must be real. The lemma implies that this real root is ± 1, as desired.

Remark. The geometric meaning of this problem is that every rotation in 3-dimensional space has a fixed line, called the *axis of rotation*. That line will be an eigenspace with eigenvalue ± 1. Our solution generalizes to show that a rotation in $\mathbb{R}^n$ admits a fixed line if n is odd. (A degree

[1] See `https://math.stackexchange.com/a/653143`

n real polynomial must have a real root if n is odd.) But a rotation in $\mathbb{R}^n$ when n is even need not have a fixed line. This is the first hint that rotation groups behave differently in even and odd dimensions.

(h) True. If P is a projection matrix, then null(P) is the space of eigenvectors with eigenvalue 0, and col(P) is the space of eigenvectors with eigenvalue 1 (because $x = Py$ for some y if and only if $x = Px$). The space col(P) is the space you are projecting onto, while null(P) is the space that is killed under the projection. These two spaces collectively span the domain of P, so we can find a basis consisting of eigenvectors of P, with eigenvalues 0 and 1. Therefore, these are the only eigenvalues of P (and P is diagonalizable).

(7) (a) $\text{rank}(B) = 2$.

(b) $\det(B^\top B) = 0$.

(c) Cannot be determined. (The eigenvalues of $B^\top B$ are the squares of the singular values of B. The singular values of B cannot be determined from the eigenvalues of B.)

(d) The eigenvalues are $\{1, \frac{1}{2}, \frac{1}{5}\}$.

(8) Let λ_i be the diagonal entries of Λ, which are positive because A is positive definite. Let $v_1, \ldots, v_n$ be the columns of Q.

The axes point in the directions specified by $v_1, \ldots, v_n$.

The semiaxis length for the axis in direction v_i is $\frac{1}{\sqrt{\lambda_i}}$.

(9) $\det(2M) = 2^{2020} \det(M)$.

(10) We have $\text{rank}(P) = \text{trace}(P)$. This follows from our solution to (6.h). Indeed, the rank of P is the dimension of col(P), which is the multiplicity of the eigenvalue 1, which is the sum of eigenvalues of P (since all eigenvalues are 0 or 1), which is the trace of P.