**Problem 0:** What is Jordan’s Canonical Form of \[
\begin{bmatrix}
2 & 1 & 0 \\
0 & 2 & 1 \\
0 & 0 & 3 \\
\end{bmatrix}
\]?

**Solution 0:**
\[
\begin{bmatrix}
2 & 1 & 0 \\
0 & 2 & 0 \\
0 & 0 & 3 \\
\end{bmatrix}
\]

**Problem 1:** Here is what is known about a linear operator \(L\) that maps a vector space to itself: No matrix \(L\) that represents \(L\) in any basis can have 0 as its first diagonal element, but this element is 3 in at least one such \(L\). What operator is \(L\)? Justify your answer.

**Solution 1:** We shall see why \(L = 3I\) where \(I\) is the identity operator. Evidently \(L \neq O\), so \(w^T L \neq o^T\) for some functional \(w^T \neq o^T\). Suppose now that a vector \(v\) existed satisfying \(w^T v = 1\) and \(w^T L v = 0\); then a basis \(B = \{v, b_2, b_3, \ldots\}\) could be chosen in which \(b_2, b_3, \ldots\) is a basis for the subspace annihilated by \(w^T\), so that every \(w^T b_j = 0\), and then \(w^T\) would be the first “row” in the inverse basis \(B^{-1}\), whereupon the matrix \(L = B^{-1}LB\) that represents \(L\) in this basis would have \(w^T L v = 0\) for its first diagonal element. This can’t happen, according to the problem’s statement. Therefore no vector \(v\) can ever satisfy both \(w^T v = 1\) and \(w^T L v = 0\); therefore every \(w^T L = \mu w^T\) for some scalar \(\mu = \mu(w^T) \neq 0\). This implies \(L = B^{-1}LB\) is diagonal for every basis \(B\). No two diagonal elements can differ without violating the equation \(w^T L = \mu w^T\) when \(w^T\) is the difference between their corresponding “rows” in \(B^{-1}\). This makes \(L\) a scalar multiple of the identity matrix \(I\), and therefore \(L\) is a scalar multiple of the identity operator \(I\); the scalar turns out to be 3.

**Problem 2:** \(Z\) is called a “Commutator” just when \(Z = XY – YX\) for some matrices \(X\) and \(Y\) (not determined uniquely by \(Z\)). Prove that if \(Z\) is a commutator, so is \(\bar{Z} = \begin{bmatrix}
0 & r^\top \\
0 & c \\
\end{bmatrix}\) for any row \(r^\top\) and column \(c\).

**Solution 2:** Suppose \(Z = XY – YX\); this equation remains valid if \(X\) is replaced by \(X + \beta I\) for any scalar \(\beta\), so we might as well assume \(X\) is invertible. Then \(\bar{Z} = \bar{X}Y – Y\bar{X}\) wherein
\[
\bar{X} = \begin{bmatrix}
0 & o^\top \\
o & X \\
\end{bmatrix}
\quad \text{and} \quad
\bar{Y} = \begin{bmatrix}
0 & -r^\top X^{-1} \\
X^{-1} c & Y \\
\end{bmatrix}.
\]
**Bonus Problem 1/2:** Prove that if \( \text{Trace}(Z) = 0 \) then \( Z \) is a commutator.

**Solution 1/2:** The asserted inference is obviously valid if \( Z \) is 1-by-1 or a bigger zero matrix. Therefore assume that \( Z \) is a nonzero matrix of dimension bigger than 1. Our proof goes by induction; we assume the desired inference valid for all matrices of dimension smaller than \( Z \)’s with \( \text{Trace} \) zero. \( Z \) cannot be a nonzero scalar multiple of the identity since \( \text{Trace}(Z) = 0 \), so Problem 1 implies that some invertible \( C \) exists such that \( C^{-1}ZC = 0 \); observe next that \( \text{Trace}(B) = 0 \). Then the induction hypothesis implies that \( B \) is a commutator, whereupon Problem 2 implies that \( C^{-1}ZC = XY - YX \) is a commutator too, whence so is \( Z \).

This proof is shorter than was found by K. Shoda (1936) *Japan J. Math.* 13 361-5, and again by A.A. Albert and B. Muckenhoupt (1957) *Michigan Math. J.* 4 1-3.

**Problem 3:** For any two vectors \( u \) and \( v \) in Euclidean 3-space \( E^3 \), an anti-commutative non-associative *Cross-Product* \( u \times v = -v \times u \) is defined as follows from the column vectors \( u \) and \( v \) that represent their respective vectors \( u = Bu \) etc. in *any* orthonormal basis \( B \):

If \( u = \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} \) and \( v = \begin{bmatrix} \xi \\ \eta \\ \zeta \end{bmatrix} \) then \( u \times v := \begin{bmatrix} \beta \xi - \gamma \eta \\ \gamma \zeta - \alpha \xi \\ \alpha \eta - \beta \zeta \end{bmatrix} \).

\( u \times v \) does depend upon basis \( B \). Changing to a new orthonormal basis \( BQ^{-1} \) where \( Q^{-1} = Q^T \) changes the respective representatives of \( u \), \( v \) and \( u \times v \) from \( u \), \( v \) and \( u \times v \) to \( Qu \), \( Qv \) and \((Qu) \times (Qv) = Q(u \times v) \cdot \text{det}(Q)\), thereby changing \( u \times v \) by a factor \( \text{det}(Q) \).

What values can \( \text{det}(Q) \) take?

Let \( X(u) \) be the linear operator acting upon \( E^3 \) that maps \( v \) to \( X(u)v = u \times v \); in that basis \( B \) some matrix \( X(u) \) represents \( X(u) \). To what does the foregoing change of basis change the operator \( X(u) \)? After that prove \( X(u)^2 = uu^T - u^TU \) regardless of orthonormal basis, and then find the eigenvalues of \( X(u) \).

**Solution 3:** In that orthonormal basis \( B \), \( X(u) \) has the matrix \( X(u) = X(u) = \begin{bmatrix} \alpha & \beta & \gamma \\ -\beta & \gamma & 0 \\ -\alpha & 0 & \gamma \end{bmatrix} := \begin{bmatrix} 0 & -\gamma & \beta \\ \gamma & 0 & -\alpha \\ -\beta & \alpha & 0 \end{bmatrix} \),

whence \( X(u)^2 = uu^T - u^TU \), but this by itself does not prove \( X(u)^2 = uu^T - u^TU \) since we have not yet established how \( X(u) \) changes when the basis changes. First we observe that \( \text{det}(Q) = \text{det}(Q^T) = \text{det}(Q^{-1}) = 1/\text{det}(Q) \) so \( \text{det}(Q) = \pm 1 \). Changing old basis \( B \) to new basis \( BQ^{-1} \) changes \( \text{old}(X(u))v = \text{old}(u \times v) \) to \( \text{new}(X(u))v = \text{new}(u \times v) = \text{old}(u \times v) \cdot \text{det}(Q) \) so that \( \text{old}(X(u)) \) changes to \( \text{new}(X(u)) = \text{old}(X(u)) \cdot \text{det}(Q) \). Therefore \( X(u)^2 = uu^T - u^TU \) for every orthonormal basis. Though Problem 3 didn’t ask, \( X(u) \) changes to \( X(Qu) = QX(u)Q^T \cdot \text{det}(Q) \).

There is another way to verify that \( X(u)^2 = uu^T - u^TU \); use the so-called *triple cross-product* formula \( w \times (u \times v) = uw^T v - vw^T u \) and set \( w = u \).
Problem 4 (continued from Problem 3): Assume what Problem 3 asked to be proved, and that $u^T u = 1$ and $u$ is a real scalar. Obtain a short formula for $\exp(\mu X(u))$ as a polynomial in $X(u)$ and $\cos(\mu)$ and $\sin(\mu)$, and then explain why $\exp(\mu X(u))$ is the operator that rotates a vector about axis $u$ through an angle $\mu$. (Recall that $\exp(i \mu) = \cos(\mu) + i \sin(\mu)$ if $1 = \sqrt{-1}$.)

Solution 4: Abbreviate $X(u) = X$, and deduce that $X^3 = -X$ as was done in the last paragraph of Solution 3. Therefore $X^{2n+1} = (-1)^n X$. Then the series for the exponential can be split into

\[
\exp(\mu X) = 1 + \sum_{n>0} \mu^n X^n / n! \\
= 1 + \sum_{n\geq0} \mu^{2n+1} X^{2n+1} / (2n+1)! + \sum_{n>0} \mu^{2n} X^{2n} / (2n)! \\
= 1 + \sum_{n\geq0} \mu^{2n+1} (-1)^n X / (2n+1)! - \sum_{n>0} \mu^{2n} (-1)^n X^2 / (2n)! \\
= 1 + X \sin(\mu) + X^2 (1 - \cos(\mu)),
\]

which exhibits the desired polynomial. Confirm that $\exp(\mu X)^{-1} = \exp(-\mu X) = \exp(\mu X)^T$ since $X^T = -X$. Therefore $\exp(\mu X)$ is an orthogonal operator that preserves Euclidean length for all $\mu$. Next choose an arbitrary nonzero $v$ and let $w(\mu) := \exp(\mu X)v$ for all real $\mu$. Now, $u^T w(\mu) = u^T v$ because $u^T X = 0^T$, and $w(\mu)^T w(\mu) = v^T v$, which implies that the cosine of the angle between $u$ and $w(\mu)$ stays the same as between $u$ and $v$ for all $\mu$. Therefore the angle between $u$ and $w(\mu)$ stays constant as $w(\mu)$ moves; it must be revolving about the axis $u$ at some rate determinable by differentiation: $d w / d \mu = X w = u x w$ stays orthogonal to both $u$ and $w(\mu)$ and has constant squared magnitude $w^T X^T X w = -v^T X^2 v$. This means $w(\mu)$ revolves with constant angular velocity performing one revolution when $\mu$ increases by $2\pi$. Therefore $\mu$ is the angle of revolution through which $\exp(\mu X(u))$ rotates any vector about the axis $u$.

Problem 5: A Simplex in an $n$-dimensional vector space is the convex hull of any $n+1$ points that do not lie in any hyperplane of dimension less than $n$. Each of those points is a vertex of the simplex. Opposite every vertex is a face of that simplex consisting of the convex hull of the other $n$ vertices. For example, in 3-space the simplex is a tetrahedron and its faces are triangles. Suppose column vectors $x_0, x_1, x_2, \ldots, x_n$ are given as the coordinates of the vertices for an orthonormal basis of an Euclidean space. Exhibit one short formula for the $n$-dimensional unoriented content (like volume) of the simplex, and another for the $(n-1)$-dimensional unoriented content (like area) of the face opposite $x_0$, whose vertices are at $x_1, x_2, \ldots, x_n$. (Unoriented contents are the nonnegative magnitudes of oriented contents.)
**Solution 5:** The unoriented content of the simplex is $|\det((x_1-x_0, x_2-x_0, x_3-x_0, \ldots, x_n-x_0)|/n!$ because it is a fraction $1/n!$ of the content of a parallelepiped from whose corner at $x_0$ radiate $n$ edges to all the other vertices of the simplex. The fraction $1/n!$ comes from repeated integrations of the cross-sections parallel to a face of each of successive simplices of lower dimensions.

Let $n$-by-$(n-1)$ matrix $F = [x_2-x_1, x_3-x_1, x_4-x_1, \ldots, x_n-x_1]$ . The $(n-1)$-dimensional unoriented content of the face opposite $x_0$ turns out to be $\sqrt{(\det(F^T F))/(n-1)!}$. To see why this is so, let $W = W^T = W^{-1}$ be the elementary orthogonal reflector that swaps the last column of the $n$-dimensional identity matrix $I$ with a unit vector orthogonal to all $n-1$ columns of $F$. This $W$ reflects the given simplex and its faces without altering their unoriented contents. The last row of $R := WF$ is a row of zeros; the columns of $R$ are the vectors emanating from $Wx_1$ to the other $n-1$ vertices of the reflected image of the face opposite $x_0$. This image is an $(n-1)$-dimensional simplex whose unoriented content is $|\det(first$ $n-1$ $rows$ $of$ $R)|/(n-1)!$. This is the same as $\sqrt{(\det(R^T R))/(n-1)!} = \sqrt{(\det(F^T F))/(n-1)!}$.

**Problem 6:** A *Polar Factorization* $F = QH$ of a given possibly rectangular real matrix $F$ has $Q^T Q = I$ and $H$ *Symmetric Positive Semidefinite*, which means that $H = H^T$ and $x^T H x \geq 0$ for every $x$. It is analogous to the representation of each complex number $f = q \cdot h$ as a product of its magnitude $h = |f|$ by a complex number $q$ of magnitude $|q| = 1$. Which real matrices $F$ have a polar factorization? Why?

**Solution 6:** If $F$ has more columns than rows it has no polar factorization because $Q$ would have to have more columns than rows and couldn’t satisfy $Q^T Q = I$ since the rank of a product $(I)$ can’t exceed the rank of a factor $(Q)$. Otherwise, when $F$ has no more columns than rows, $F = QH$ obtainable as follows: Let $R$ be an orthogonal matrix of eigenvectors of $F^T F = RV^2 R^T$ so that $R^T = R^{-1}$ and $V^2$ is a diagonal matrix of eigenvalues of $F^T F$. We know that $V^2 \geq 0$ elementwise because $x^T V^2 x = (FRx)^T (FRx) \geq 0$ for every $x$. Therefore diagonal matrix $V \geq 0$ can be obtained from $V^2$ by taking square roots elementwise. Then $H := RVR^T$ must satisfy $F^T F = H^2$; in other words, $H$ is the symmetric positive semidefinite square root of $F^T F$.

Life is easy when the columns of $F$ are linearly independent. Then $x^T V^2 x = (FRx)^T (FRx) > 0$ for every $x \neq 0$, so that all diagonal elements of $V$ are positive; and $Q = FH^{-1} = FRV^{-1} R^T$. When the columns of $F$ are linearly dependent, some diagonal element(s) of $V$ must vanish, and then the corresponding columns of $FR$ must vanish too because $(FR)^T (FR) = V^2$. Dividing the nonzero columns of $FR$ by the corresponding nonzero diagonal elements of $V$ produces a matrix whose nonzero columns are orthonormal; replace its zero columns by more orthonormal columns, as must be possible because the total number of orthonormal columns will not exceed the number of rows, to obtain finally a matrix $P$ of orthonormal columns that satisfies $P^T P = I$ and $PV = FR$. Now set $Q := PR^T$ to find $Q^T Q = P^T RR^T P = I$ and $QH = PR^T RVR^T = F$.

Incidentally, if non-square $F$ has a polar factorization, $F^T$ doesn’t, and *vice-versa.*