This is a CLOSED BOOK EXAM to which you may bring NO textbooks and ONE sheet of notes. Solve problem 1 first, and then as many subsequent problems as you can, in three hours. If a problem seems too hard, come back to it after trying another. Complete solutions earn far more credit than partial or defective solutions, so take your time and take pains to get them right. On each page you hand in for a grade you must put your name and the number(s) of the problem(s) solved thereon to get credit.

1: What makes matrix notation so valuable, namely the ability to represent so many things by matrices, can also make it confusing; the notation does not say what the matrix represents. A real matrix L may represent, among other things, …

- **•0:** a change of coordinates (basis) in a real vector space,
- **•1:** a linear map **L** from one real vector space to another,
- **•2:** a linear map **L** from one Euclidean vector space to another,
- **•3:** a linear map **L** of a real space to itself,
- **•4:** a linear map **L** from a real vector space to its dual space,
- **•5:** a quadratic form £(**x**) for vectors **x** in a non-Euclidean real space, or
- **•6:** a quadratic form £(**x**) for vectors **x** in an Euclidean space.

Changes of coordinates in the relevant space(s) affect L differently in each case; for each case explain how, and describe (without proof) as many as you can of the attributes of L unaltered by such changes.

Solution: Let the nonsingular matrices E and F represent changes of coordinates, and let $\overline{L}$ be the matrix that represents, after changes of coordinates, whatever L represented before.

1•0: If **B** is one basis for a vector space, **B**L a second, and (**B**L)E = **B**$\overline{L}$ a third, it could have been reached directly from the first basis by using $\overline{L}$:= LE . All that L and $\overline{L}$ need have in common are their dimension and the nonvanishing of their determinants.

1•1: If **L** maps a vector space with basis **B** to another space with basis **C** then **L** is represented by the matrix L := $\mathbf{C}^{-1}\mathbf{LB}$ because **C**y = **y** = **Lx** = **LB**x for column-vectors x and y that satisfy y = Lx . Changing bases from **B** to **B**E and from **C** to **C**F changes L to $\overline{L} := F^{-1}LE$. This is an *Equivalence*; it has to preserve only dimensions and rank.

1•2: If **L** maps one Euclidean space with orthonormal basis **B** to another with orthonormal basis **C** , changes to new orthonormal bases **B**E and **C**F that preserve the root-sum-squares formula for length must be accomplished by orthogonal matrices: $E^T = E^{-1}$ and $F^T = F^{-1}$. Such changes of bases change the matrix L := $\mathbf{C}^{-1}\mathbf{LB}$ that represents $\mathbf{C}^{-1}\mathbf{LB}$ to $\overline{L} := F^{-1}LE$. This is an *Orthogonal Equivalence*; it has to preserve only dimensions and singular values.

1•3: If **L** maps a vector space with basis **B** to itself, **L** is represented by matrix L := $\mathbf{B}^{-1}\mathbf{LB}$. Changing to a new basis **B**E changes L to $\overline{L} := E^{-1}LE$. This *Similarity* need preserve only the Jordan Normal form in which the order of Jordan blocks is immaterial. Because L is real, complex eigenvalues and eigenvectors, if any, come in complex conjugate pairs which can be exhibited without any complex arithmetic by the real Jordan Normal form with 2-by-2 real blocks on the diagonal instead of 1-by-1 complex eigenvalues.

1•4: If **L** maps a real vector space with basis **B** to its dual, where the dual basis is $\mathbf{B}^{-1}$, the scalar-value of the bilinear form **Lxy** , acting linearly upon each vector $\mathbf{x} = \mathbf{B}x$ and $\mathbf{y} = \mathbf{B}y$, is obtained from their column vectors of coordinates and a matrix L that represents **L** thus: $\mathbf{Lxy} = (Lx)^T y$. To obtain L given **L** , substitute various unit column-vectors for x and y . Changing to a new basis **B**E changes x to $\bar{x} := E^{-1}x$, y to $\bar{y} := E^{-1}y$, and L to $\bar{L} := E^T L E$, so that $\mathbf{Lxy} = (\bar{L}\,\bar{x})^T \bar{y}$ too. This *Congruence* relation between L and $\bar{L}$ is an equivalence, so it preserves the dimension and rank of L as well as the *Signature* of its symmetric part $(L^T+L)/2$; see 1•5.

1•5: The quadratic form £(**x**) is obtained, for vectors $\mathbf{x} = \mathbf{B}x$ in a real space with basis **B** , from the column x that represents **x** and a matrix L that represents £ thus: $£(\mathbf{x}) = x^T L x$. Only the symmetric part $(L^T+L)/2$ of L matters here, so we might as well assume they are the same. To obtain L given £ , proceed as in 1•4 from the symmetric bilinear form defined by $\mathbf{Lxy} := \big(£(\mathbf{x}+\mathbf{y}) - £(\mathbf{x}-\mathbf{y})\big)/4$. Changing to a new basis **B**E changes L to $\bar{L} := E^T L E$ and preserves its dimension, symmetry and signature (the numbers of L's positive, negative and zero eigenvalues — see 1•6).

1•6: Continuing from 1•5, if the space is Euclidean the change from one orthonormal basis **B** to another, **B**E , requires that E be orthogonal: $E^T = E^{-1}$. Now the congruence $\bar{L} = E^T L E$ is an *Orthogonal Similarity* that preserves the eigenvalues, all real, of symmetric matrix L .

==

2: Suppose that $F = LCR^T$ in which L, C and R all have the same rank and C is square and invertible. Explain why $L^\dagger = (L^T L)^{-1} L^T$, $R^\dagger = (R^T R)^{-1} R^T$ and $F^\dagger = R^{\dagger T} C^{-1} L^\dagger$.

Explanation: Not " $(A \cdot B)^\dagger = B^\dagger A^\dagger$ ", which is often false; try $((uv^T)(wx^T))^\dagger$. The formulas for $L^\dagger$ and $R^\dagger$ follow from the observation that L, C and R all have the same number of columns, and that number is their rank, so $(\ldots)^{-1}$'s exist. Let $G := R^{\dagger T} C^{-1} L^\dagger$. Since C and $R^\dagger R = L^\dagger L = I$ have the same dimensions, $GF = R^{\dagger T} C^{-1} L^\dagger L C R^T = R^{\dagger T} R^T = R R^\dagger = (GF)^T$. Similarly $FG = LL^\dagger = (FG)^T$. Then $FGF = LL^\dagger L C R^T = F$, and similarly $GFG = G$. Therefore G satisfies all four equations that determine the Moore-Penrose Pseudo-inverse $F^\dagger$ of F uniquely.

==

3: Given constants μ_0, μ_1, μ_2 and $\mu_3 \neq 0$ set $x_{n+1} := \mu_0 x_n + \mu_1 x_{n-1} + \mu_2 x_{n-2} + \mu_3 x_{n-3}$ for $n = 3, 4, 5, \ldots$ in turn. Show for all these n that, regardless of x_0, x_1, x_2 and x_3,

$$\det\left(\begin{bmatrix} x_{n+3} & x_{n+2} & x_{n+1} & x_n \\ x_{n+2} & x_{n+1} & x_n & x_{n-1} \\ x_{n+1} & x_n & x_{n-1} & x_{n-2} \\ x_n & x_{n-1} & x_{n-2} & x_{n-3} \end{bmatrix}\right)/(-\mu_3)^n \quad \text{is independent of } n .$$

Proof: Let X_{n+3} be the matrix in the determinant, and set $M := \begin{bmatrix} \mu_0 & 1 & 0 & 0 \\ \mu_1 & 0 & 1 & 0 \\ \mu_2 & 0 & 0 & 1 \\ \mu_3 & 0 & 0 & 0 \end{bmatrix}$. Then it is easy to verify that $X_{n+3} = X_{n+2} \cdot M = X_3 \cdot M^n$. Therefore $\det(X_{n+3}) = \det(X_3) \cdot \det(M^n) = \det(X_3) \cdot (-\mu_3)^n$.

==

4: An *Unreduced Upper-Hessenberg Matrix* is a square matrix with no zeros on the first sub-diagonal and zeros everywhere below it. The *Jordan-Normal Form* of a *Derogatory* matrix has at least two different *Jordan-Blocks* with the same eigenvalue. Are there any derogatory unreduced Hessenberg matrices? Justify your answer.

Answer: No. Consider any derogatory N-by-N matrix B . It must have at least one eigenvalue ß for which B–ßI (which has the same rank as the Jordan Normal form to which it is *Similar*) has rank less than N–1 because two or more Jordan Blocks of B–ßI have zeros on their diagonals. But the first sub-diagonal of an unreduced N-by-N upper Hessenberg matrix H is the diagonal of an (N–1)-by-(N–1) submatrix with nonzero determinant, which implies $\mathrm{rank}(H - ßI) \geq N-1$ for every scalar ß .

==

5: In a vector space, a set is called "Connected" if every two of its members are joined by some continuous path consisting entirely of members of that set. In the space of all real N-by-N matrices the orthogonal matrices do not form a connected set; prove this, and also prove that the N-by-N complex unitary matrices do form a connected set.

Proof: Since det(…) is a continuous function of its argument, and since every real orthogonal matrix has determinant +1 or else –1 , the two kinds of orthogonal matrices cannot form one connected set.

Now let $P = P^{*-1}$ be any N-by-N unitary matrix; it has a Schur decomposition $P = QUQ^*$ in which Q is unitary and U is upper-triangular. However, it is easy to verify that $U^* = U^{-1}$ must be simultaneously upper- and lower-triangular, and is therefore a diagonal matrix which can be written $U = \exp(\imath H)$ for some real diagonal H and $\imath = \sqrt{(-1)}$. For $0 \leq \mu \leq 1$ set $P(\mu) := Q\exp(\imath\mu H)Q^*$ to describe a continuous path from $P(0) = I$ to $P(1) = P$. Given any two unitary N-by-N matrices B and C , set $P := B^*C$ to obtain a continuous path $BP(\mu)$ from $B = BP(0)$ to $C = BP(1)$. End of proof.

(More generally, the invertible matrices form a connected subset among N-by-N complex matrices, but not among real. Can you see why?)

6: The *Schur Complement* of the scalar ß in the complex square matrix $B := \begin{bmatrix} \beta & r^* \\ c & F \end{bmatrix}$ is $E := F - cr^*/ß$. Here r^* is the complex conjugate transpose of r . Now, assuming that $\mathrm{Re}\{ z^*Bz/z^*z \} \ge \mu > 0$ for every nonzero complex vector z of the appropriate dimension, prove that the same is true with E in place of B . (Hint: Try a real non-symmetric matrix B and real z first.) What does that assumed inequality imply about the diagonal elements of the upper-triangular factor U in the factorization $B = L\cdot U$ with a unit-lower-triangular L ?

Proof: For any scalar π and nonzero vector z of the appropriate dimension, define

$f(\pi, z) := [\pi^*, z^*]B \begin{bmatrix} \pi \\ z \end{bmatrix} /(|\pi|^2 + z^*z) = (ß|\pi|^2 + \pi^*r^*z + \pi z^*c + z^*Fz)/(|\pi|^2 + z^*z)$, noting that

$\mathrm{Re}\{f\} \ge \mu$ no matter how π and z are chosen. In particular $\mathrm{Re}\{ß\} \ge \mu > 0$. Consequently $z^*Ez/z^*z = (z^*Fz - z^*cr^*z/ß)/z^*z = (1 + |\pi|^2/z^*z)f(\pi, z) - (ß\pi + r^*z)(ß\pi^* + z^*c)/(ßz^*z)$. Now set $\pi := -r^*z/ß$ or $-(z^*c/ß)^*$ to find then $z^*Ez/z^*z = (1 + |\pi|^2/z^*z)f(\pi, z)$, so its real part can be no less than μ , as claimed. This inequality implies the same inequality for every diagonal element of the upper-triangular factor U , the first of which is ß and the rest are the upper-left corner elements of successive Schur complements. Compare the class notes on "Diagonal Prominence".

==

7: P and Q are *Orthogonal Projectors* from an Euclidean Space into itself; this means that $P^T = P = P^2$ and $Q^T = Q = Q^2$. The norm $\|\ldots\|$ is the biggest singular value. Prove that …

a) $\|P-Q\| \le 1$.

b) If $\mathrm{Rank}(P) \ne \mathrm{Rank}(Q)$ then $\|P-Q\| = 1$, but not conversely.

c) If $\|P-Q\| < 1$ then Rank(P) = Rank(Q) and
no nonzero vector in Range(Q) is orthogonal to Range(P) .

d) The converse of (c). (This is harder.)

Proof 7(a): Since P–Q is real symmetric, its singular values are the magnitudes of its eigenvalues. All the eigenvalues of positive semidefinite P are zeros or ones, and likewise for Q ; therefore no eigenvalue of P–Q can exceed 1 , nor fall below –1 . Therefore $\|P-Q\| \le 1$.

7(b): If, say, Rank(P) > Rank(Q) , then Range(P) and Nullspace(Q) must have a nonzero intersection since their dimensions add up to more than the dimension of the space; therefore $Qx = o \ne Px = x$ for some x in that intersection, and $1\cdot\|x\| \ge \|P-Q\|\cdot\|x\| \ge \|(P-Q)x\| = \|x\| \ne 0$ and consequently $\|P-Q\| = 1$. The converse is untrue because $\|P-Q\| = 1$ for projectors

$P = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $Q = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ that have the same rank 1 .

Proof 7(c): Let $q = Qq \ne o$ be any nonzero vector in Range(Q) . Now, $Pq = P^Tq = o$ if q is orthogonal to Range(P) , and then $\|P-Q\|\cdot\|q\| \ge \|(P-Q)q\| = \|q\| \ne 0$ so $\|P-Q\| = 1$; therefore if $\|P-Q\| < 1$ no nonzero q in Range(Q) can be orthogonal to range(P) , and vice-versa (swapping Q and P), and Rank(P) = Rank(Q) because of part (b).

7(d): Conversely, suppose $\|P-Q\| = 1$. Then $(P-Q)x = \pm x \ne o$ for at least one eigenvector x . There are now two cases to consider: In the first case $(P-Q)x = -x$ and then $Px = o$ and $x = Qx \ne o$ because $0 \le \|Px\|^2 = x^TPx = x^TQx - x^Tx \le x^Tx - x^Tx = 0$; this x must be in Range(Q) and orthogonal to Range(P) . In the second case $(P-Q)x = +x$ and then similarly $Qx = o$ and $x = Px \ne o$; this means that Range(P) and Nullspace(Q) have a nonzero intersection. This can occur if Range(Q) is a proper subspace of Range(P) , in which case Rank(P) > Rank(Q) , and then no nonzero vector in Range(Q) need be orthogonal to Range(P) . But otherwise, when r := Rank(P) = Rank(Q) too in the second case, then also n := Nullity(P) = Nullity(Q) , and then

$$\begin{aligned} 0 &= n+r - \mathrm{Dim}(\mathrm{Range}(P)) - \mathrm{Dim}(\mathrm{Nullspace}(Q)) \quad \ldots \quad (\, n+r = \mathrm{Dim}(\text{whole space})\,) \\ &< n+r - \mathrm{Dim}(\mathrm{Range}(P)) - \mathrm{Dim}(\mathrm{Nullspace}(Q)) + \mathrm{Dim}(\,\mathrm{Range}(P) \cap \mathrm{Nullspace}(Q)\,) \\ &= n+r - \mathrm{Dim}(\,\mathrm{Range}(P) + \mathrm{Nullspace}(Q)\,) = \mathrm{Dim}(\,(\mathrm{Range}(P) + \mathrm{Nullspace}(Q))^{\perp}\,) \\ &= \mathrm{Dim}(\,\mathrm{Range}(P)^{\perp} \cap \mathrm{Nullspace}(Q)^{\perp}\,) = \mathrm{Dim}(\,\mathrm{Nullspace}(P) \cap \mathrm{Range}(Q)\,)\ . \end{aligned}$$

Therefore some vector in Range(Q) is orthogonal to Range(P) , as claimed.