This is an OPEN BOOK EXAM to which you may bring any papers and textbooks. Solve as many 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.

Problem 0: Give concrete numerical examples of three systems Ax = b of linear equations, one system under-determined, a second system over-determined, and a third both over- and under-determined, in which all elements of A and b are nonzero. (" Over-determined " means some of the system's equations are redundant. " Under-determined " means solutions are not unique.)

Solution 0: A = [1, 1] and b = [1] for under-determined. A = b = $\begin{bmatrix}1\\1\end{bmatrix}$ for over-determined.
A = $\begin{bmatrix}1&1\\1&1\end{bmatrix}$ and b = $\begin{bmatrix}1\\1\end{bmatrix}$ for both over- and under-determined.

Problem 1: Acting upon the linear space of polynomials of degree at most 3 , the linear operator **D** maps a polynomial to its derivative. What is Jordan's Normal Form of **D** ?

Solution 1: These polynomials form a 4-dimensional space.
Since $\mathbf{D}^4 = O \ne \mathbf{D}^3$, Jordan's Normal Form of **D** is $\begin{bmatrix}0&1&0&0\\0&0&1&0\\0&0&0&1\\0&0&0&0\end{bmatrix}$.

Problem 2: **K** is a linear operator. If **KX** is a basis for Range(**K**) and **Z** is a basis for Nullspace(**K**) , then [**X**, **Z**] is a basis for Domain(**K**) . Prove it.

Solution 2: Check dimensions first:

Dimension(Range(**X**)) ≥ dimension(Range(**KX**)) = dimension(Range(**K**))

since **KX** is a basis for Range(**K**) . Since

dimension(Range(**K**)) + dimension(Nullspace(**K**)) = dimension(Domain(**K**)) ,

the number of (column) vectors in [**X**, **Z**] is at least dimension(Domain(**K**)) . All that is left to do to prove that [**X**, **Z**] is a basis is to show that the vectors in it are linearly independent. So suppose **Xu** + **Zv** = **o** ; then **KXu** + **KZv** = **KXu** = **o** too since **Z** is a basis for Nullspace(**K**) . But **KX** is a basis, so **u** = **o** . Therefore **Zv** = **o** , which means **v** = **o** too. Therefore the vectors in [**X**, **Z**] are independent; they form a basis for Domain(**K**) .

Problem 3: Prove *Ptolemy's Inequality* : $\|z-x\|\cdot\|y\| \le \|y-x\|\cdot\|z\| + \|z-y\|\cdot\|x\|$ for any three nonzero vectors x, y, z in an Euclidean space, with equality just when $x/\|x\|^2$, $y/\|y\|^2$ and $z/\|z\|^2$ lie in that order on a straight line.

Solution 3: Into the Triangle Inequality $\|\bar{z}-\bar{x}\| \le \|\bar{y}-\bar{x}\| + \|\bar{z}-\bar{y}\|$ for $\bar{x} := x/\|x\|^2$ etc., substitute

$$\|\bar{y} - \bar{x}\|^2 = 1/\|x\|^2 - 2x^Ty/(\|x\|\cdot\|y\|)^2 + 1/\|y\|^2 = (\|y-x\|/(\|x\|\cdot\|y\|))^2 \text{ , etc.}$$

Equality in the triangle inequality occurs just when $\bar{x}$, $\bar{y}$, $\bar{z}$ lie in that order on a straight line.

Problem 4: Given 2x2 real matrix $L = \begin{bmatrix} \lambda & 0 \\ \rho & \mu \end{bmatrix}$ exhibit a real orthogonal Q such that $LQ = QL^T$.

Solution 4: $Q = \dfrac{\begin{bmatrix} \lambda-\mu & \rho \\ \rho & \mu-\lambda \end{bmatrix}}{\sqrt{(\lambda-\mu)^2+\rho^2}} = Q^T = Q^{-1}$ unless $\rho = \lambda-\mu = 0$, in which case set $Q := I$.
(The first column of Q must be an eigenvector of L .)

Problem 5: Suppose E is the matrix of a linear operator that maps one Euclidean space to another in such a way as preserves angles; *i.e.*, $\cos(\angle(Ex, Ey)) = \cos(\angle(x, y)) := x^Ty/(\|x\|\cdot\|y\|)$ for all nonzero column vectors x and y . Must E have orthogonal columns all of the same Euclidean length? Explain why.

Solution 5: Yes, and here is why: Set $M := E^TE = M^T$; then all nonzero x and y must satisfy $(y^TMx)^2/(y^TMy\ x^TMx) = (y^Tx)^2/(y^Ty\ x^Tx)$. Therefore $y^TMx = 0$ if and only if $y^Tx = 0$. Since this is true for every y^T , one of Fredholm's Alternatives implies that $Mx = \mu x$ for some scalar μ which may, for all we know so far, vary with x . Since this is true for every x , every nonzero x is an eigenvector of M , which implies that $M = \mu I$. Therefore $E^TE = \mu I$, and $\mu > 0$ because $0 < \|Ex\|^2 = x^TE^TEx = \mu x^Tx$, so E 's columns are orthogonal with length $\sqrt{\mu}$.

Alternative solution 5: Let u_j be the j-th column of the identity matrix so that Eu_j is the j-th column of E . Eu_j is orthogonal to Eu_k whenever $j \ne k$ because u_j is orthogonal to u_k . And Eu_j+Eu_k is orthogonal to Eu_j-Eu_k for a similar reason; this makes $\|Eu_j\| = \|Eu_k\|$.

Problem 6: Linear operator $\mathbf{L}$ maps one Euclidean vector space to another; what is the maximum value taken by $\mathbf{u}^T\mathbf{L}\mathbf{v}$ as $\mathbf{u}$ and $\mathbf{v}$ run over all unit-vectors ($\mathbf{u}^T\mathbf{u} = \mathbf{v}^T\mathbf{v} = 1$) ?

Solution 6: The maximum of $\mathbf{u}^T\mathbf{L}\mathbf{v}$ over all unit-vectors u and v is the biggest singular value μ of $\mathbf{L}$. This is so because orthonormal coordinate systems can be chosen in the domain and target spaces of $\mathbf{L}$ that represent it by a diagonal matrix L of singular values, and then $(u^TLv)^2 \le u^Tu(Lv)^TLv$ (by Cauchy's inequality — see the previous problem) is maximized when u and v each has just one nonzero component (1) in a location corresponding to the biggest singular value(s) μ .

Alternative solution 6: Assume $\mathbf{L} \ne O$ lest the problem be trivial. Let $\mathbf{H} := \begin{bmatrix} O & L \\ L^T & O \end{bmatrix}$. By definition, the nonzero singular values of $\mathbf{L}$ are the positive eigenvalues of $\mathbf{H}$. An orthonormal basis can be chosen to represent $\mathbf{H}$ by a diagonal matrix of $\mathbf{H}$'s eigenvalues; then for every $\mathbf{y} \ne \mathbf{o}$ we find that $\mathbf{y}^T\mathbf{H}\mathbf{y}/\mathbf{y}^T\mathbf{y}$ is a weighted average of those eigenvalues maximized when $\mathbf{y}$ has one nonzero component for the biggest eigenvalue, and then this maximum $\mathbf{y}^T\mathbf{H}\mathbf{y}/\mathbf{y}^T\mathbf{y}$ is the biggest singular value of $\mathbf{L}$. Let $\mathbf{y}^T = [\mathbf{u}^T, \mathbf{v}^T]$ and $\mu := \mathbf{y}^T\mathbf{H}\mathbf{y}/\mathbf{y}^T\mathbf{y} = 2\mathbf{u}^T\mathbf{L}\mathbf{v}/(\mathbf{u}^T\mathbf{u} + \mathbf{v}^T\mathbf{v})$ be the maximizing vector and maximized value. Replacing $\mathbf{u}$ by $\mathbf{u}/\sqrt{\mathbf{u}^T\mathbf{u}}$ and $\mathbf{v}$ by $\mathbf{v}/\sqrt{\mathbf{v}^T\mathbf{v}}$ increases μ unless $\mathbf{u}^T\mathbf{u} = \mathbf{v}^T\mathbf{v}$, so this must already hold. Again, $\max \mathbf{u}^T\mathbf{L}\mathbf{v} = \mu$.

Problem 7: The 3-by-3 nilpotent matrix N satisfies $N^3 = O \neq N^2$. The set of all matrices P that *commute* with N (they satisfy NP = PN) constitutes a subspace in the 9-dimensional space of 3-by-3 matrices; what is the subspace's dimension, and why?

Solution 7: The subspace's dimension is 3 and here is why: Choose a coordinate system that exhibits N in its Jordan Normal Form, which is one 3-by-3 Jordan block. Then it is easy to show that P must be a quadratic polynomial in N by solving NP = PN for P element by element from lower left by diagonals to upper right in that coordinate system.

Problem 8: Explain why the identity $\begin{bmatrix} CB & O \\ B & O \end{bmatrix} = \begin{bmatrix} I & C \\ O & I \end{bmatrix} \begin{bmatrix} O & O \\ B & BC \end{bmatrix} \begin{bmatrix} I & -C \\ O & I \end{bmatrix}$ implies that the matrix products BC and CB , if both exist, have the same *nonzero* eigenvalues. If B and C are square too, BC and CB have the same eigenvalues; then must BC and CB be *Similar* too? Say why. Try simple examples like B = [1, 0] and $C = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ before jumping to conclusions.

Solution 8: If both exist, BC and CB must be square though perhaps of different dimensions. The identity is a *Similarity* because $\begin{bmatrix} I & C \\ O & I \end{bmatrix}^{-1} = \begin{bmatrix} I & -C \\ O & I \end{bmatrix}$, so $\begin{bmatrix} CB & O \\ B & O \end{bmatrix}$ and $\begin{bmatrix} O & O \\ B & BC \end{bmatrix}$ have the same eigenvalues.These eigenvalues are zeros and the eigenvalues of CB , or of BC ; therefore CB and BC have the same nonzero eigenvalues. But they are not necessarily similar, not even if they have the same dimensions. For example take $B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ and $C = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ to find $BC = B \neq O$ but CB = O . (However, if B and C are square and either is invertible then BC *is* Similar to $CB = B^{-1}(BC)B$.)

Problem 9: Suppose H and W–H are real symmetric positive definite matrices; why must $H^{-1} - W^{-1}$ be positive definite too?

Solution 9: Since H is positive definite, there is a congruence that diagonalizes W and H simultaneously; say $H = C^{T-1}IC^{-1}$ and $W = C^{T-1}VC^{-1}$ for some diagonal matrix V . (For instance, the Choleski factorization of $H = U^TU$ and the eigensystem factorization of $U^{T-1}WU^{-1} = QVQ^T$ with $Q^T = Q^{-1}$ provide $C = U^{-1}Q$.) Since $W–H = C^{T-1}(V–I)C^{-1}$ is positive definite, so is its congruent V–I , which means every diagonal element of V exceeds 1 . Now $H^{-1} - W^{-1} = C(I - V^{-1})C^T$ is positive definite because its congruent $I - V^{-1}$ is.

Alternative solution 9: Identity $H^{-1} - W^{-1} = W^{-1}(W–H)W^{-1} + W^{-1}(W–H)H^{-1}(W–H)W^{-1}$ expresses $H^{-1} - W^{-1}$ as a sum of positive definite matrices and thus positive definite. Another identity $H^{-1} - W^{-1} = H^{-1}(H^{-1} + (W–H)^{-1})^{-1}H^{-1}$ works too but is harder to derive.

Problem 10: Given matrices F, g, b and C, none of them square, show that a solution-pair {u, n} of the equations $F^TFu - C^Tn = F^Tg$ and $Cu = b$, if any solution-pair exists, must minimize the sum of squares $(Fu - g)^T(Fu - g)$ over all u constrained by $Cu = b$. If this constraint is not inconsistent, why must a finite solution-pair {u, n} always exist?

Solution 10: Suppose {û, ñ} is a solution-pair; it satisfies $F^T(F\hat{u}-g) = C^T\tilde{n}$ and $C\hat{u} = b$. Then any other u satisfying the constraint $Cu = b = C\hat{u}$ must have

$$\begin{aligned}(Fu-g)^T(Fu-g) - (F\hat{u}-g)^T(F\hat{u}-g) &= (F(u-\hat{u}) + F\hat{u}-g)^T(F(u-\hat{u}) + F\hat{u}-g) - (F\hat{u}-g)T(F\hat{u}-g)\\ &= (F(u-\hat{u}))^TF(u-\hat{u}) + 2(F\hat{u}-g)^TF(u-\hat{u})\\ &= (F(u-\hat{u}))^TF(u-\hat{u}) + 2\tilde{n}^TC(u-\hat{u})\\ &= (F(u-\hat{u}))^TF(u-\hat{u}) \ \geq\ 0\ .\end{aligned}$$

Therefore the minimum value of $(Fu-g)^T(Fu-g)$ subject to the constraint $Cu = b$ is achieved when $u = \hat{u}$.

According to Fredholm's Alternatives, the constraint $Cu = b$ is satisfiable (not inconsistent) if and only if $v^Tb = 0$ whenever $v^TC = o^T$. Analogously, the equation $\begin{bmatrix} F^TF & -C^T \\ C & O \end{bmatrix}\begin{bmatrix} u \\ n \end{bmatrix} = \begin{bmatrix} F^Tg \\ b \end{bmatrix}$ is satisfiable if and only if $w^TF^Tg + v^Tb = 0$ whenever $\begin{bmatrix} w^T & v^T \end{bmatrix}\begin{bmatrix} F^TF & -C^T \\ C & O \end{bmatrix} = \begin{bmatrix} o^T & o^T \end{bmatrix}$. This last equation implies $w^TF^TF - v^TC = o^T$ and $w^TC = o^T$, which implies $w^TF^TFw = v^TCw = 0$, which implies $w^TF^T = o^T$, which implies $v^TC = o^T$, which implies $v^Tb = 0$ when the constraint is satisfiable, which implies $w^TF^Tg + v^Tb = 0$ and therefore finite solution-pairs {u, n} must exist.

(This solution is closely analogous to the solution of the unconstrained Least-Squares problem.)

–•–

The exam was deemed long enough that the next problem was not needed, but it is included here just for the record.

Problem -1: A skew-symmetric bilinear operator $\mathbf{W}$ is defined for any linear functional $\mathbf{w}^T \neq \mathbf{o}^T$ thus: $\mathbf{Wxy} := \mathbf{xw}^T\mathbf{y} - \mathbf{yw}^T\mathbf{x} = -\mathbf{Wyx}$. How does the Range of $\mathbf{W}$ compare with the Nullspace of $\mathbf{w}^T$, and why?

Solution -1: They are the same; here is why: Evidently $\mathbf{w}^T\mathbf{Wxy} = 0$ for all $\mathbf{x}$ and $\mathbf{y}$, so the Range of $\mathbf{W}$ is contained in the Nullspace of $\mathbf{w}^T$. On the other hand, for every $\mathbf{z}$ in the Nullspace of $\mathbf{w}^T$ (so that $\mathbf{w}^T\mathbf{z} = 0$), and for any $\mathbf{v}$ such that $\mathbf{w}^T\mathbf{v} \neq 0$ (such a $\mathbf{v}$ must exist because $\mathbf{w}^T \neq \mathbf{o}^T$), set $\mathbf{u} := \mathbf{v}/\mathbf{w}^T\mathbf{v}$ to infer that $\mathbf{Wzu} = \mathbf{z}$ and hence the Nullspace of $\mathbf{w}^T$ is contained in the Range of $\mathbf{W}$. Therefore the two subspaces are the same.