STAT 512 - Assignment 1 - due date is on course outline

1. Prove: If W is a vector subspace of a vector space V, then $\dim(W) \leq \dim(V)$.

2. If $\mathbf{A}$ is an $n \times m$ matrix, then $vec(\mathbf{A})$ is the $mn \times 1$ vector formed by stringing the columns of $\mathbf{A}$ out, one after the other.

   (a) Show that $tr\left(\mathbf{A}'\mathbf{A}\right) = \|vec(\mathbf{A})\|^2$. The square root of this quantity is called the *Frobenius norm* of $\mathbf{A}$, written $\|\mathbf{A}\|$.

   (b) Show that $\|\mathbf{Ax}\| \leq \|\mathbf{A}\| \, \|\mathbf{x}\|$.

   (c) A vector-matrix equation $\mathbf{Ax} = \mathbf{w}$ might be solved for $\mathbf{x}$ by writing $\mathbf{B} = \mathbf{I} - \mathbf{A}$ and solving $\left(\mathbf{I} - \mathbf{B}\right)\mathbf{x} = \mathbf{w}$ by iteratively calculating

   $$\mathbf{x}_n = \mathbf{B}\mathbf{x}_{n-1} + \mathbf{w}.$$

   Show that $\|\mathbf{x}_{n+1} - \mathbf{x}_n\| \leq \|\mathbf{B}\| \, \|\mathbf{x}_n - \mathbf{x}_{n-1}\|$. If $\|\mathbf{B}\| < 1$ then the iterative scheme is called a *contraction*, and we will later see that it is guaranteed to converge to a solution.

3. Show that a matrix is non-negative definite if and only if it is the covariance matrix of some random vector.

4. Prove: Any symmetric matrix can be represented as the difference between two p.s.d. matrices which are mutually orthogonal and the sum of whose ranks equals the rank of the original matrix.

5. Show that the diagonal elements of a symmetric matrix are bounded above and below by the maximum and minimum eigenvalues, respectively. Conclude that a matrix with both positive and negative diagonal elements cannot be positive definite.

6. Prove (Theorem 2.3.9 in Khuri): If $\mathbf{A}$ is an $m \times n$ matrix and $\mathbf{B}$ is a $n \times m$ matrix $(m \leq n)$, then the eigenvalues of $\mathbf{BA}$ are those of $\mathbf{AB}$ together with $n - m$ zeros.

7. Prove (Theorem 2.3.17 in Khuri): If $\mathbf{A}$ is symmetric and $\mathbf{B} > \mathbf{0}$, both $n \times n$, then $\mathbf{x}'\mathbf{Ax}/\mathbf{x}'\mathbf{Bx}$ lies between the minimum and maximum eigenvalues of $\mathbf{B}^{-1}\mathbf{A}$.

...over

8. A 'leading principal minor' of a square matrix is the determinant of the submatrix formed from the first m rows and columns.

   (a) The LU-decomposition: Suppose that a matrix $\mathbf{M}_{n\times n}$ is such that *all* leading principal minors are non-zero. Show that we can represent $\mathbf{M}$ as $\mathbf{M} = \mathbf{LU}$, where $\mathbf{L}$ and $\mathbf{U}$ are nonsingular, with $\mathbf{L}$ lower triangular and $\mathbf{U}$ upper triangular (both nonsingular).

   (b) Suppose $\mathbf{M}_{n\times n}$ is symmetric and positive definite. Show that we can represent $\mathbf{M}$ as $\mathbf{M} = \mathbf{W}'\mathbf{W}$ with $\mathbf{W}$ upper triangular. This is called the 'Cholesky decomposition' of a p.d. matrix.

9. Let $\mathbf{x}$ be a random vector with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Let $\mathbf{A}$ be a symmetric matrix. Show - without making any mention of the individual elements - that
$$E\left[\mathbf{x}'\mathbf{A}\mathbf{x}\right] = tr\left(\mathbf{A}\boldsymbol{\Sigma}\right) + \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}.$$

10. Let $\mathbf{X}$ be an $n \times p$ matrix with rank $r \leq \min(n, p)$.

    (a) Show that $\mathbf{X}$ can be factored as $\mathbf{X} = \mathbf{FG}$, where $\mathbf{F}_{n\times r}$ and $\mathbf{G}_{r\times p}$ both have rank r.

    (b) Define $\mathbf{X}^+ = \mathbf{G}'(\mathbf{G}\mathbf{G}')^{-1}(\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}' : p\times n$. This matrix $\mathbf{X}^+$ is the *Moore-Penrose generalized inverse* of $\mathbf{X}$. Verify (to your own satisfaction – I don't need to see it) that

    $\mathbf{X}\mathbf{X}^+\mathbf{X} = \mathbf{X}$, $\mathbf{X}^+\mathbf{X}\mathbf{X}^+ = \mathbf{X}^+$, $\mathbf{X}\mathbf{X}^+$ and $\mathbf{X}^+\mathbf{X}$ are symmetric and idempotent.

    (c) Show that the set of solutions to the equations $\mathbf{X}\boldsymbol{\theta} = \mathbf{y}$ (assuming that there is a solution) is the set of vectors $\boldsymbol{\theta} = \mathbf{X}^+\mathbf{y} + \mathbf{c}$, where $\mathbf{X}\mathbf{c} = \mathbf{0}$, and that the shortest (i.e. minimum norm) such solution is $\mathbf{X}^+\mathbf{y}$.

(Hint for (c): Let $\boldsymbol{\theta}_0$ be one solution. This gives you useful information about the structure of $\mathbf{y}$.)

Note: If there is no solution to the equations in (c) then $\hat{\boldsymbol{\theta}} = \mathbf{X}^+\mathbf{y}$ is still the <u>minimizer</u> of $\|\mathbf{y} - \mathbf{X}\boldsymbol{\theta}\|$, and $\textsc{cov}\left[\hat{\boldsymbol{\theta}}\right] = \sigma^2\mathbf{X}^+\mathbf{X}^{+\prime}$, which (by verifying the properties above) turns out to $= \sigma^2\left(\mathbf{X}'\mathbf{X}\right)^+$. Example: To get the least squares estimates in regression, we seek $\hat{\boldsymbol{\theta}}$ minimizing $\|\mathbf{y} - \mathbf{X}\boldsymbol{\theta}\|$. If $\mathbf{X}$ has full rank p, then in the above $\mathbf{F} = \mathbf{X}$ and $\mathbf{G} = \mathbf{I}_p$, so that $\mathbf{X}^+ = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$. In this case, that the minimizer of $\|\mathbf{y} - \mathbf{X}\boldsymbol{\theta}\|$ is $\hat{\boldsymbol{\theta}} = \mathbf{X}^+\mathbf{y}$ was shown in Lecture 4. If $\mathbf{X}$ has less than full rank then the minimizer is not unique (and $\mathbf{X}^+ \neq (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$), but $\hat{\boldsymbol{\theta}} = \mathbf{X}^+\mathbf{y}$ is still *a* minimizer, and is the *shortest minimizer*. At this point you are able to prove all of these statements, although that is not being asked.

...over

Addendum (to finish off the proof of the Spectral Decomposition Theorem)

Result: If λ_0 is an eigenvalue with multiplicity r, of a real symmetric matrix $\mathbf{M}_{n\times n}$, then the vector space $\mathcal{V}$ of eigenvectors corresponding to λ_0 has dimension r.

Proof: Define $\mathbf{M}_0 = \mathbf{M} - \lambda_0\mathbf{I}$, and let $s = rk\,(\mathbf{M}_0)$. We first show that $s = n - r$. For this, write the characteristic polynomial of $\mathbf{M}$ as

$$|\mathbf{M} - \lambda\mathbf{I}| = (\lambda - \lambda_0)^r\, p(\lambda),$$

where $p(\lambda_0) \neq 0$. Then the characteristic polynomial of $\mathbf{M}_0$ is

$$|\mathbf{M}_0 - \lambda\mathbf{I}| = |\mathbf{M} - (\lambda + \lambda_0)\,\mathbf{I}| = \lambda^r p(\lambda + \lambda_0),$$

with exactly r roots equal to zero since $p(\lambda_0) \neq 0$. Thus $\mathbf{M}_0$ has $n - r$ zero eigenvalues and r nonzero eigenvalues. As in Question #10 of this assignment, we can write $\mathbf{M}_0 = \mathbf{FG}$, where $\mathbf{F}_{n\times s}$ and $\mathbf{G}_{s\times n}$ are each of rank s. Then since

$$s = rk\mathbf{M}_0 = rk\mathbf{M}_0\mathbf{M}_0' = rk\mathbf{M}_0^2 = rk\mathbf{FGFG} \leq rk\,(\mathbf{GF})_{s\times s} \leq s,$$

we conclude that $\mathbf{GF}$ is nonsingular and hence all s of its eigenvalues are nonzero. But as in Question #7, the $n - r$ nonzero eigenvalues of $\mathbf{FG}$ $(= \mathbf{M}_0)$ coincide with the s nonzero eigenvalues of $\mathbf{GF}$, so that $s = n - r$.

Now with $\mathbf{F}_{n\times n-r}$ and $\mathbf{G}_{n-r\times n}$ as above, (note that $\mathbf{F}$ has rank $n-r$ and so $\mathbf{F}'\mathbf{F}$ is invertible) we have that

$$\mathcal{V} = \{\mathbf{v} \mid \mathbf{Mv} = \lambda_0\mathbf{v}\} = \{\mathbf{v} \mid \mathbf{M}_0\mathbf{v} = \mathbf{0}\} = \{\mathbf{v} \mid \mathbf{Gv} = \mathbf{0}\} = \mathrm{col}\left(\mathbf{I}_n - \mathbf{G}'\,(\mathbf{GG}')^{-1}\,\mathbf{G}\right),$$

with

$$\dim(\mathcal{V}) = rk\left(\mathbf{I}_n - \mathbf{G}'\,(\mathbf{GG}')^{-1}\,\mathbf{G}\right) = tr\left(\mathbf{I}_n - \mathbf{G}'\,(\mathbf{GG}')^{-1}\,\mathbf{G}\right) = n - (n - r) = r,$$

as required. □

Added note: We have used here the fact that if $\mathbf{H}_{n\times n}$ is idempotent, then its rank equals its trace. But the proof of this from class used the result that we are currently proving, and so we should prove it in another way here. Let r be the rank of $\mathbf{H}$, and write (using #10 again) $\mathbf{H} = \mathbf{AB}$, where $\mathbf{A}_{n\times r}$ and $\mathbf{B}_{r\times n}$ are each of rank r. Then since $\mathbf{H}$ is idempotent, we have

$$\mathbf{A}\,(\mathbf{I}_r - \mathbf{BA})\,\mathbf{B} = \mathbf{0},$$

and since $\mathbf{A}'\mathbf{A}$ and $\mathbf{BB}'$ are invertible, we conclude that $\mathbf{BA} = \mathbf{I}_r$. Thus

$$tr\mathbf{H} = tr\mathbf{AB} = tr\mathbf{BA} = r.$$