This is the front page of the practice exam. When you come to the real exam, you'll see here a place to write your name and student number. The duration of the final exam is 180 minutes. The final exam is of similar question composition and format to this practice exam.

Instructions:

There are 5 questions, with 20 points each. Each question has two items. Item (a) worth 14 points and item (b) worth 6 points.

Answer all questions on this exam paper.

Use the other side of the paper if additional space is required.

Question 1

(a) Consider the data matrix,

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 1 \end{bmatrix},$$

with Singular Value Decomposition (SVD) of the form

$$A = U \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} V^T,$$

where U is a 4×2 matrix with orthonormal columns, and V is a 2×2 matrix with orthonormal columns. Determine the singular values σ_1 and σ_2.

Solution: The singular values are the square roots of the positive eigenvalues of $A^T A$.

$$A^T A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}.$$

And,

$$\begin{vmatrix} 2-\lambda & 1 \\ 1 & 1-\lambda \end{vmatrix} = (2-\lambda)(1-\lambda) - 1 = \lambda^2 - 3\lambda + 1.$$

Hence,

$$\lambda_{1,2} = \frac{3 \pm \sqrt{9-4}}{2} = 2.618,\ 0.382.$$

Hence $\sigma_1 = \sqrt{2.618} = 1.618$ and $\sigma_2 = \sqrt{0.382} = 0.618$.

(b) Consider now the matrix $B = A\,A^T\,A$. What are the singular values of B?

Solution: Use the the SVD of $A = U\Sigma V^T$.

$$A\,A^T\,A = U\Sigma V^T V\Sigma U^T U\Sigma V^T = U\Sigma^3 V^T = U \begin{bmatrix} \sigma_1^3 & 0 \\ 0 & \sigma_2^3 \end{bmatrix} V^T.$$

Hence the singular values are $\sigma_1^3 = 4.235$ and $\sigma_2^3 = 0.236$.

Question 2

(a) Assume you are presented with tuples of data points $(x_{11}, x_{21}, G_1, y_1), \ldots, (x_{1n}, x_{2n}, G_n, y_n)$. Where x's and y's are real values and $G_i \in \{\text{'m'}, \text{'f'}\}$ where 'm' stands for 'male' and 'f' stands for 'female'. To describe this data, you wish to fit a model,

$$y = \exp\left(\begin{cases} \beta_m + \beta_1 x_1 + \beta_2 x_2 & \text{if male,} \\ \beta_f + \beta_1 x_1 + \beta_2 x_2 & \text{if female.} \end{cases}\right).$$

You do this by selecting $\beta = [\beta_0, \beta_1, \beta_2, \beta_3]^T$ that will minimize,

$$L = \sum_{i=1}^{n} \big(\log y_i - \beta_0 - \beta_1 x_{1i} - \beta_2 x_{2i} - \beta_3 I_i \big)^2,$$

where $I_i = 0$ if $G_i =$ ''m' and $I_i = 1$ if $G_i =$ 'f'. You then set $\beta_m = \beta_0$ and $\beta_f = \beta_0 + \beta_3$.
To minimize L, you set $\hat{\beta} = (A^T A)^{-1} A^T v$ where $A \in \mathbb{R}^{n \times 4}$ and $v \in \mathbb{R}^n$. Assume that n is even and the data is sorted such that the first $1, \ldots, n/2$ observations are 'male' and the others 'female'. Determine A and v.

Solution: This is simply setting up a data-fitting problem using least squares.

$$A = \begin{bmatrix} 1 & x_{11} & x_{21} & 0 \\ 1 & x_{12} & x_{22} & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{1,n/2} & x_{2n/2} & 0 \\ 1 & x_{1,n/2+1} & x_{2,n/2+1} & 1 \\ 1 & x_{1,n/2+2} & x_{2,n/2+2} & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{1n} & x_{2n} & 1 \end{bmatrix}, \qquad v = \begin{bmatrix} \log y_1 \\ \log y_2 \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \log y_n \end{bmatrix}.$$

(b) Say that when you attempt to compute $(A^TA)^{-1}$ you find out that A does not have linearly independent columns and thus A^TA is singular. You then resort to ridge regression by aiming to minimize,

$$||A\beta - v||^2 + \lambda||\beta||^2.$$

with $\lambda > 0$. The minimizer is given by $\hat{\beta}_\lambda = (A^TA + \lambda I)^{-1}A^Tv$. Does the inverse in this formula exist for any $\lambda > 0$ or are there cases where it doesn't? Prove your claim.

Solution: This problem is a least squares problem aiming to minimize $||\tilde{A}\beta - v||^2$ with,

$$\tilde{A} = \begin{bmatrix} A \\ \sqrt{\lambda} I \end{bmatrix}.$$

First observe that $\tilde{A}$ has linearly independent columns. For this assume $\tilde{A}z = 0$ and you need to show that $z = 0$. However $\tilde{A}z = 0$ implies that $\sqrt{\lambda}Iz = 0$ or $Iz = 0$ or $z = 0$. Hence $\tilde{A}$ has linearly independent columns.

Now observe that the Gram matrix of $\tilde{A}$ is $\tilde{A}^T\tilde{A} = A^TA + \lambda I$. It is known that the Gram matrix of a matrix with linearly independent columns is non-singular.

(proof of last statement - not needed for solution): Assume $\tilde{A}^T\tilde{A}w = 0$. We want to show that $w = 0$ and hence the columns of $\tilde{A}^T\tilde{A}$ are linearly independent and hence the same with the rows and hence it is non-singular. To show this, left multiply by w^T to get.

$$w^T\tilde{A}^T\tilde{A}w = 0.$$

Or $||\tilde{A}w||^2 = 0$. Or $||\tilde{A}w|| = 0$. Hence by the properties of a norm, $\tilde{A}w = 0$. Now the assumption implies that $w = 0$.

Question 3

(a) Consider the vectors $v_1 = [1, 0, 1]^T$, $v_2 = [6, 6, 0]^T$ and the 3×2 matrix $A = [v_1\ v_2]$. Determine a QR factorization of A having the form $A = QR$ where Q is a 3×2 matrix with orthonormal columns and R is an upper triangular matrix. Throughout this question, avoid using decimal points, but rather use exact arithmetic.

Solution: $\tilde{q}_1 = v_1$. $||\tilde{q}_1|| = \sqrt{2}$. Hence $q_1 = [\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}]^T$.

$\tilde{q}_2 = v_2 - (v_2^T q_1)q_1 = [6, 6, 0] - 3\sqrt{2}[\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}]^T = [3, 6, -3]^T$.

$||\tilde{q}_2|| = \sqrt{54} = 3\sqrt{6}$. Hence $q_2 = [\frac{1}{\sqrt{6}}, \frac{\sqrt{6}}{3}, -\frac{1}{\sqrt{6}}]^T$.
Now

$$Q = [q_1\ q_2].$$

To get R, use $R_{11} = ||\tilde{q}_1|| = \sqrt{2}$, $R_{22} = ||\tilde{q}_2|| = 3\sqrt{6}$ and $R_{12} = q_1^T v_2 = 3\sqrt{2}$.

In summary $A = QR$ is:

$$\begin{bmatrix} 1 & 6 \\ 0 & 6 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{\sqrt{6}}{3}, \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} \sqrt{2} & 3\sqrt{2} \\ 0 & 3\sqrt{6} \end{bmatrix}.$$

(b) Assume you now don't know the values in A and only have the values in R that you computed above. Use the values of R to determine the eigenvalues of the matrix AA^T.

Solution: We remember that the non-zero eigenvalues of AA^T are the same as those of A^TA. So let's look at A^TA. Since $A = QR$ we have $A^TA = R^TQ^TQR = R^TR$. Now we need to compute R^TR:

$$R^TR = \begin{bmatrix} 2 & 6 \\ 6 & 72 \end{bmatrix}.$$

Hence the characteristic polynomial equated to zero is $(2-\lambda)(72-\lambda) - 36 = 0$. Or,

$$\lambda^2 - 74\lambda + 108 = 0$$

This means,

$$\lambda_{1,2} = \frac{74 \pm \sqrt{74^2 - 4 \times 108}}{2} = 37 \pm \sqrt{1261}.$$

Hence $\lambda_1 = 72.51$ and $\lambda_2 = 1.489$.

Now AA^T is of dimension 3×3 and thus an additional eigenvalue must be $\lambda_3 = 0$ since AA^T cannot be full rank.

Question 4

(a) Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$, given by $f(x) = (x - d)^T C\,(x - d) + b^T x$ with,

$$C = \begin{bmatrix} 4 & 2 \\ 2 & -2 \end{bmatrix}, \qquad b = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad \text{and} \qquad d = \begin{bmatrix} 6 \\ 4 \end{bmatrix}.$$

Compute the gradient function $\nabla f(x)$ and determine a point $x \in \mathbb{R}^2$ for which x, $\nabla f(x) = 0$. Is this a local minimum point?

(b) Consider now a least squares problem, minimizing $||Ax - b||$ where $A \in \mathbb{R}^{n\times 2}$ and $b \in \mathbb{R}^n$. Observe the following computation (completion of the square),

$$\begin{aligned} ||Ax - b||^2 &= (Ax - b)^T(Ax - b) \\ &= x^TA^TAx - 2b^TAx + b^Tb \\ &= x^TA^TAx - 2b^TAx + b^Tb - 2b^TA^TAx + 2b^TA^TAx \\ &= x^TA^TAx - 2b^TA^TAx + b^Tb + 2b^TA^TAx - 2b^TAx \\ &= (x - b)^TA^TA(x - b) + 2b^T(A^TA - A)x. \end{aligned}$$

Is there a matrix A and vector b such that the least square problem is equivalent to minimizing $f(\cdot)$ from (a)? If so, express A and b in terms of C, D and d, otherwise, explain why there isn't such an A.

Question 5

(a) Consider the matrices,

$$A = \begin{bmatrix} 1/2 & 1 \\ 0 & 2 \end{bmatrix}, \qquad \text{and} \qquad B = \begin{bmatrix} 2 & -1 \\ 0 & 1/2 \end{bmatrix}.$$

Determine the eigenvalues of the matrix $B^{20}A^{22}$.

Solution: A thing to recognize here is that $B = A^{-1}$ (or alt. $A = B^{-1}$). Hence $B^{20}A^{22} = A^2$.

The eigenvalues of A are $1/2$ and 2. This pops out immediately for a diagonal matrix since the determinant is the product of the diagonal entries. Hence the eigenvalues of A^2 are the eigenvalues of A each squared (this comes directly from the eigenvalue/eigenvector equation $Ax = \lambda x$. Hence the eigenvalues are $1/4$ and 4.

(b) Consider now the 4×4 matrix W,

$$W = \begin{bmatrix} B^{20}A^{22} & B^{20}A^{7} \\ 0 & B^{20}A^{19} \end{bmatrix}$$

Determine the eigenvalues of W.

Solution: First using the fact that $A = B^{-1}$ obtain,

$$W = \begin{bmatrix} A^{2} & B^{13} \\ 0 & B \end{bmatrix}.$$

Now you can see that the eigenvalues of a block upper diagonal matrix are the eigenvalues of the block diagonal components. One way to show this is via $\det(W - \lambda I_4) = \det(A^2 - \lambda I_2)\det(B - \lambda I_2)$ - a fact not discussed in class much - as we didn't do much on determinants.

Hence the eigenvalues of W are $1/4$ and 4 (coming from A^2) and 2 and $1/2$ (coming from B).

END OF EXAMINATION