NAME: Preston Chan !!

503943014

CS170A — Mathematical Models and Methods for Computer Science

Spring 2013

D.S. Parker, Scott G-H. Tu

Midterm Examination

OPEN BOOK, OPEN NOTES, CLOSED NETWORK

Wednesday, May 8, 4:00pm–5:50pm

Do not cheat.

Problem	*Points*
1	14 /25
2	15 /25
3	18 /25
4	23 /25
Total	70 /100

Throughout this exam we will make use of the following dataset D of height and weight values:

i	h_i	w_i	
1	58	115	2
2	59	117	
3	60	120	3
4	61	123	3
5	62	126	3
6	63	129	3
7	64	132	3
8	65	135	3

0 2 5 8 11 14 17 20

7

15

26

40

57

77

$\frac{72}{8}$ 115 8 $\frac{5}{8}$

123

Let $\boldsymbol{h}$ be the column vector of h_i values, and $\boldsymbol{w}$ be the column vector of w_i values.
For the column vectors $\boldsymbol{h}$ and $\boldsymbol{w}$, the least squares fit to $\boldsymbol{w} = a\boldsymbol{h} + b$ is $a = 2.917$, $b = -54.75$.

61.5

2.917
· 61.5
14585
29170
1750200
160435.5
− 54.78

$\det = \frac{1}{2}$

1. SVD

For each of the following matrices, determine (yes/no) whether the matrix is: unitary, hermitian, invertible, positive definite.

matrix	unitary?	hermitian?	invertible?	positive definite?
$\frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}$	True ☐ False ☒	True ☐ False ☒	True ☒ False ☐	True ☐ False ☒
$\frac{1}{\sqrt{2}}\begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}$	True ☐ False ☒	True ☒ False ☐	True ☒ False ☐	True ☐ False ☒
$\frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$	True ☐ False ☒	True ☐ False ☒	True ☒ False ☐	True ☐ False ☒

For this examination, assume:

Definition: A is *positive definite* if it is Hermitian and all its eigenvalues are positive real values.

For each of the following equations, mark whether the equation is True (valid for all specified matrices A) or False. Assume that A' denotes the hermitian transpose of A, and $i = \sqrt{-1}$.

For full credit, if you mark it True, you must *explain* how you derived this. If you mark it False, you must give a *counterexample*.

(a) **True** ☒ **False** ☐

If X and Y are real matrices such that $X + iY$ is unitary, then $\begin{pmatrix} X & -Y \\ Y & X \end{pmatrix}$ is orthogonal.

Explanation: unitary $X + iY \Rightarrow$ X unitary, Y unitary, X real, Y real $\Rightarrow \begin{pmatrix} X & -Y \\ Y & X \end{pmatrix}$ orthogonal

(b) **True** ☐ **False** ☒

If both the determinant and trace of a 2×2 Hermitian matrix A are positive real values, then A is positive definite.
(Hint: the trace of A — the sum of its diagonal elements — is also the sum of its eigenvalues.)

Explanation: $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ eigenvalues $3, -1$ ∴ not positive definite

$\det(A) = 1$

$tr(A) = 1$

(c) **True** ☒ **False** ☐

If H is Hermitian, then $(I + iH)(I - iH)^{-1}$ is unitary, where I is the identity.
(Hint: $(A^{-1})' = (A')^{-1}$).

Explanation:

$I + iH$ is hermitian

$I - iH = (I + iH)'$

$$((I+iH)(I-iH)^{-1})' \stackrel{?}{=} ((I+iH)(I-iH)^{-1})^{-1}$$

$$((I-iH)^{-1})'(I+iH)' \stackrel{?}{=} ((I-iH)^{-1})^{-1}(I+iH)^{-1}$$

$$((I-iH)')^{-1}(I+iH)' = (I-iH)(I+iH)^{-1}$$

(d) **True** ☒ **False** ☐

If C is a correlation matrix with SVD $C = U S V'$, then $U = V$ and S is a diagonal matrix with positive values on the diagonal.

Explanation: The SVD of C is identical to its Eigendecomposition, since C is real and symmetric. Therefore the SVD is in the form of an eigendecomposition: $C = USV' = QLQ'$, $U = V$
(S has positive values on diag since S is singular values $= |\lambda_i|$

Covariance

Assume D is the 8×2 matrix with the two columns $\boldsymbol{h}$ and $\boldsymbol{w}$ mentioned earlier.

(a) What differences are there between the covariance matrices for D and for the dataset that is like D, but has two identical copies of every row?

$$\mathrm{cov}(D) = \frac{2(n-1)}{2n-1}\,\mathrm{cov}(\text{new } D)$$

new D $\quad$ D

$$\frac{2}{2n-1}(x-m)'(x-m) \qquad \frac{1}{n-1}(x-m)'(x-m)$$

(b) Suppose that we changed the vector $\boldsymbol{h}$ by subtracting its mean $\bar{h} = mean(\boldsymbol{h})$ from each of its entries. How would cov(D) change, if at all?

Not at all. The mean for h would just be 0, resulting in the same value when doing the matrix multiplication.

(c) Suppose that we changed the vector $\boldsymbol{h}$ by multiplying each of its entries by $1/\sigma$, where σ is the standard deviation of $\boldsymbol{h}$. How would cov(D) change, if at all?

Cov(D)

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \to \begin{pmatrix} \frac{a}{6^2} & \frac{b}{6} \\ \frac{c}{6} & d \end{pmatrix}$$

AA All

(d) Suppose that we changed $\boldsymbol{w}$ to be $\boldsymbol{h}$. How would cov(D) change, if at all?

All values in cov(D) would be the same.

-5

(e) If we plot a least squares line fit through the (h, w) points, will the point made up of the mean values $(\bar{h}, \bar{w})$ also be on the line?

-5 $\quad \bar{h} = 61.5 \quad \bar{w} = 120\frac{5}{8}$

No

$w = 2.917 \cdot 61.5 - 54.75$

$= 160.343 - 54.75$

$= 105.68$

3. **Least Squares**

As stated earlier, for the column vectors $\boldsymbol{h}$ and $\boldsymbol{w}$, the least squares fit to $\boldsymbol{w} = a\boldsymbol{h} + b$ is $a = 2.917$, $b = -54.75$.

Let '**1**' be the 8×1 vector of 1's, and let X be the matrix with the two columns $\boldsymbol{h}$ and **1**.

(a) Suppose we use least squares to solve

Σh_i^2

$$X \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \boldsymbol{w}.$$

$\begin{pmatrix} h & \cdots \\ 1 & \cdots \end{pmatrix}\begin{pmatrix} h & 1 \\ \vdots & \vdots \\ & 1 \end{pmatrix}$

What must be the resulting values for α and β?

$X'X \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = X'w$

$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = (X'X)^{-1} X'w$

$= \frac{1}{8\Sigma h_i^2 - (\Sigma h_i)^2}\begin{pmatrix} 8\Sigma h_i w_i - \Sigma h_i \Sigma w_i \\ -\Sigma h_i w_i \Sigma h_i + \Sigma h_i^2 \Sigma w_i \end{pmatrix}$

$X'X = \begin{pmatrix} \Sigma h_i^2 & \Sigma h_i \\ \Sigma h_i & 8 \end{pmatrix}$

$(X'X)^{-1} = \frac{1}{8\Sigma h_i^2 - (\Sigma h_i)^2}\begin{pmatrix} 8 & -\Sigma h_i \\ -\Sigma h_i & \Sigma h_i^2 \end{pmatrix}$

$X'w = \begin{pmatrix} h & \cdots \\ 1 & \end{pmatrix}\begin{pmatrix} w \\ \vdots \end{pmatrix} = \begin{pmatrix} \Sigma h_i w_i \\ \Sigma w_i \end{pmatrix}$

(b) Is is true that $||\boldsymbol{w}||^2 = (n-1)\,\mathrm{var}(\boldsymbol{w}) + n\,(\mathrm{mean}(\boldsymbol{w}))^2$, where $n = 8$?

Yes

(c) Estimate the R^2 value for this least squares fit.

Close to 1

(d) Is it possible for pseudoinverse X^- here to be roughly the following matrix?

(T) + goes to $\emptyset$

minimizes large h values

$$\begin{pmatrix} -0.08 & -0.06 & -0.04 & -0.01 & 0.01 & 0.04 & 0.06 & 0.08 \\ 5.25 & 3.79 & 2.32 & 0.86 & -0.61 & -2.07 & -3.54 & -5.00 \end{pmatrix}.$$

proper dimensions

$h \quad 1$

Yes

(e) Give a matrix A for which $\boldsymbol{c} = A^- \boldsymbol{w}$ is the set of coefficients $\begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}$ of a quadratic polynomial fitting the dataset D using least squares.

7 - 5

PCA

The covariance matrices for each of the four matrices D is:

$$\text{cov}(D) = \begin{pmatrix} 6.0 & 17.50 \\ 17.5 & 51.13 \end{pmatrix}$$

$A - \lambda I$ $\begin{pmatrix} 6 - \lambda & 17.5 \\ 17.5 & 51.13 - \lambda \end{pmatrix}$

$\lambda^2 - 57.13\lambda +$

and each has the SVD

$$\text{cov}(D) = \begin{pmatrix} -0.32 & -0.95 \\ -0.95 & 0.32 \end{pmatrix} \begin{pmatrix} 57.12 & 0 \\ 0 & 0.0088 \end{pmatrix} \begin{pmatrix} -0.32 & -0.95 \\ -0.95 & 0.32 \end{pmatrix}$$

(a) What is the largest eigenvalue of cov(D)?
What is the first principal component of D?

1st PC = $\begin{pmatrix} -0.32 \\ -0.95 \end{pmatrix}$

largest eigenvalue = 57.12

↑ SVD of cov == eigendecomp of cov

(b) Suppose that we changed the vector $\boldsymbol{h}$ by subtracting its mean from each of its entries. How would the first principal component of D change, if at all? 4

Center around 0

No change

(c) Suppose that we changed the vectors $\boldsymbol{h}$ and $\boldsymbol{w}$ by multiplying each of their entries by 2. How would the first principal component of D change, if at all?

No change.

(d) Suppose that we switched $\boldsymbol{w}$ and $\boldsymbol{h}$, so that the columns of D are reordered. How would the first principal component of D change, if at all?

1st PC = $\begin{pmatrix} -0.95 \\ -0.32 \end{pmatrix}$ & flipped

(e) Suppose we added a new point $(h_9, w_9) = (60, 100000)$ to D. How would the principal components for resulting 9 × 2 dataset D change? −2

The 2nd PC would become more important (

larger 2nd eigenvalue