CS 189 HW5 Write-Up

Qingyang Zhao

TOTAL POINTS

134 / 164

QUESTION 1

1 Getting Started 4 / 4

✓ + 4 pts Correct

+ 0 pts Blank/wrong information

QUESTION 2

Properties of Convex Functions 40 pts

2.1 a 10 / 10

✓ + 10 pts Correct

+ 8 pts Almost Correct

+ 5 pts Halfway Progress

+ 2 pts Trivial Progress

+ 0 pts Wrong/Missing

2.2 b 7 / 10

+ 10 pts Correct

✓ + 7 pts Uses A^(T)A instead of 2A^(T)A or doesn't simplify A^(T)A + AA^(T)but the proof follows correctly

+ 6 pts Computes the hessian correctly but doesn't correctly prove that it is PSD

+ 2 pts Correctly computes gradients or makes simple error computing hessian.

+ 0 pts Incorrect

2.3 c 10 / 10

✓ + 10 pts Correct (proof by contradiction or other)

+ 8 pts Minor logical error

+ 5 pts Reasonable approach, but incomplete

+ 2 pts Some relevant work

+ 0 pts Incorrect/Blank

2.4 d 10 / 10

✓ + 2.5 pts part 1 correct

✓ + 2.5 pts part 2 correct

✓ + 2.5 pts part 3 correct

✓ + 2.5 pts part 4 correct

+ 0 pts no part correct

QUESTION 3

Canonical Correlation Analysis 60 pts

3.1 a 10 / 10

✓ + 10 pts Correct

+ 0 pts Insufficient justification - Did not use SVD, made unjustified logical jumps

3.2 b 10 / 10

✓ + 10 pts correct

+ 8 pts almost correct

+ 5 pts halfway

+ 2 pts minimal effort

+ 0 pts no effort

3.3 c 10 / 10

✓ + 10 pts correct (doesn't have to use hint2)

+ 9 pts almost correct

+ 7 pts realize the application of hint1

+ 5 pts realize the application of vector

+ 3 pts say things about norm vector

+ 0 pts incorrect

3.4 d 10 / 10

✓ + 10 pts Correct/Partially Correct

+ 0 pts Incorrect

3.5 e 10 / 10

✓ + 10 pts Correct

+ 8 pts Almost Correct

+ 6 pts Significant Progress

+ 4 pts Some Progress

+ 2 pts Trivial Progress

+ 0 pts Blank or Incorrect

3.6 f 10 / 10

✓ + 10 pts Correct argument about uselessness of CCA and correct modification of CCA

+ 8 pts Minor error in one part

+ **5 pts** Only one part correct (or both partially correct)

+ **2 pts** Only one part partially correct

+ **0 pts** Incorrect or blank

QUESTION 4

Mooney Reconstruction 50 pts

4.1 a 3 / 10

✓ + **3 pts** **Explain how to estimate covariance matrices**

+ **3 pts** Covariance Matrices in Code

+ **2 pts** Standardized data AND subtracted mean

+ **2 pts** Scaled at the end

+ **0 pts** Wrong/Missing

4.2 b 0 / 10

✓ + **0 pts** **Blank/Wrong**

+ **2 pts** Minimal Progress

+ **5 pts** Halfway

+ **8 pts** Minor Error

+ **10 pts** Correct

4.3 c 10 / 10

✓ + **10 pts** **Correct or Minor Error**

+ **8 pts** Image is off

+ **5 pts** Partial

+ **2 pts** Reasonable attempt

+ **0 pts** Incorrect or Missing

4.4 d 0 / 10

+ **10 pts** Correct

+ **5 pts** An attempt, but plot is off

✓ + **0 pts** **No Answer / Incorrect**

4.5 e 10 / 10

✓ + **10 pts** **Complete**

+ **8 pts** Close to complete

+ **5 pts** Partial

+ **2 pts** Reasonable Attempt

+ **0 pts** Wrong or Missing

QUESTION 5

5 Fruits! 0 / 0

✓ + **0 pts** **Fruits!**

QUESTION 6

6 Your Own Question 10 / 10

✓ + **10 pts** **Correct**

+ **5 pts** Question, no solution

+ **0 pts** No question or no effort

H06-1

a)

Names	Email Address
Wan	Jhun0324@berkeley.edu

Description of Team: Best Group Ever

How did I work?

Comments:

b) I certify that all solutions are entirely in my words and that I have not looked at another student's solutions. I have credited all external sources in this write up.

Qingyang Zhao

1 Getting Started 4 / 4

✓ + 4 pts Correct

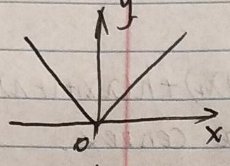
+ 0 pts Blank/wrong information

HW05-2

a) $y = f(x)$ $\quad y = |x|$ $\quad 0 \le \lambda \le 1$

$$\lambda|x_1| + (1-\lambda)|x_2| \geq |\lambda x_1 + (1-\lambda)x_2|$$

$$\parallel$$



$$|\lambda x_1| + |(1-\lambda)x_2| \geq |\lambda x_1 + (1-\lambda)x_2|$$

$\therefore$ $|x|$ is convex

but $|x|$ is not differentiable at $x=0$,

b) $f(x) = x^T A^T A x - 2y^T A x + y^T y$

$$\frac{\partial^2 f}{\partial x^2} = A^T A$$

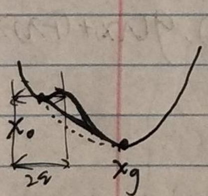
Examine : $\xi^T A^T A \xi \geq 0$

for $\forall \xi$, $\quad$ since $\|\xi^T A\|_2^2 = \xi^T A^T A \xi \geq 0$

Thus $A^T A$ is semi-definite positive.

Then Hessian of $f(x)$ is positive semi-definite

c)



① Say, there is ~~another~~ a global optimal x_g, for every x in Domain X $f(x_g) \le f(x)$

$f(x_g) < f(x_0)$

According to the definition of convex function,

$(1-\lambda)f(x_0) + \lambda f(x_g) \geq f(\lambda x_g + (1-\lambda)x_0)$ $\quad \lambda \in [0,1)$

$f(x_0) > (1-\lambda)f(x_0) + \lambda f(x_g)$

which means we can always find a λ, $\lambda x_g + (1-\lambda)x_0 \in (x_0 - \varepsilon, x_0 + \varepsilon)$

(Because, it is a line segment between x_0, x_g)

s.t. $f(x_0) > f(\lambda x_g + (1-\lambda)x_0)$ contradict to x_0 is a local minimal

which means x_0 is a global minimal.

② if $f(x_g) = f(x_0)$, then $f(x_0) \geq f(\lambda x_g + (1-\lambda)x_0)$

but $f(x_0) \le f(\lambda x_g + (1-\lambda)x_0)$

$f(x_0) = f(\lambda x_g + (1-\lambda)x_0)$ which means x_0 is a global minimal

2.1 a 10 / 10

✓ + 10 pts Correct

+ 8 pts Almost Correct

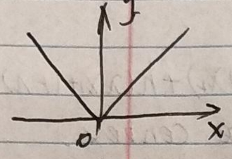
+ 5 pts Halfway Progress

+ 2 pts Trivial Progress

+ 0 pts Wrong/Missing

HW05-2

a) $y=f(x)$ $\quad y=|x|$ $\quad 0\le\lambda\le 1$



$$\lambda|x_1|+(1-\lambda)|x_2| \geq |\lambda x_1+(1-\lambda)x_2|$$
$$\parallel$$
$$|\lambda x_1|+|(1-\lambda)x_2| \geq |\lambda x_1+(1-\lambda)x_2|$$

$\therefore$ $|x|$ is convex

but $|x|$ is not differentiable at $x=0$,

b) $f(x)=x^TA^TAx-2y^TAx+y^Ty$

$$\frac{\partial^2 f}{\partial x^2}=A^TA$$

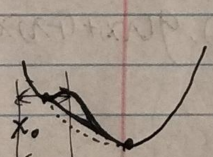
Examine: $\xi^TA^TA\xi\geq 0$

for $\forall\xi$, since $\|\xi^TA\|_2^2=\xi^TA^TA\xi\geq 0$

Thus A^TA is semi-definite positive.

Then Hessian of $f(x)$ is positive semi-definite

c)



① Say, there is global optimal x_g, for every x in Domain X $f(x_g)\le f(x)$

$f(x_g) \le f(x_0)$

According to the definition of convex function,

$(1-\lambda)f(x_0)+\lambda f(x_g)\geq f(\lambda x_g+(1-\lambda)x_0)$ $\quad\lambda\in(0,1)$

$f(x_0) \geq (1-\lambda)f(x_0)+\lambda f(x_g)$

which means we can always find a λ, $\lambda x_g+(1-\lambda)x_0\in(x_0-\epsilon,x_0+\epsilon)$

(Because, it is a line segment between x_0, x_g)

s.t. $f(x_0) \geq f(\lambda x_g+(1-\lambda)x_0)$ contradict to x_0 is a local minimal

which means x_0 is a global minimal.

② if $f(x_g)=f(x_0)$, then $f(x_0)\geq f(\lambda x_g+(1-\lambda)x_0)$

but $f(x_0)\le f(\lambda x_g+(1-\lambda)x_0)$

$f(x_0)=f(\lambda x_g+(1-\lambda)x_0)$ which means x_0 is a global minimal

2.2 b 7 / 10

+ **10 pts** Correct

✓ + **7 pts** Uses A^(T)A instead of 2A^(T)A or doesn't simplify A^(T)A + AA^(T)but the proof follows correctly

+ **6 pts** Computes the hessian correctly but doesn't correctly prove that it is PSD

+ **2 pts** Correctly computes gradients or makes simple error computing hessian.

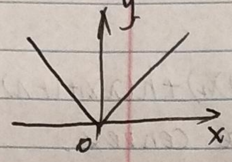
+ **0 pts** Incorrect

HW05-2

a) $y=f(x)$ $\quad y=|x|$ $\quad 0\le\lambda\le 1$

$\lambda|x_1|+(1-\lambda)|x_2| \geq |\lambda x_1+(1-\lambda)x_2|$

$\parallel$



$|\lambda x_1|+|(1-\lambda)x_2| \geq |\lambda x_1+(1-\lambda)x_2|$

$\therefore$ $|x|$ is convex

but $|x|$ is not differentiable at $x=0$,

b) $f(x)=x^TA^TAx-2y^TAx+y^Ty$

$\frac{\partial^2 f}{\partial x^2}=A^TA$

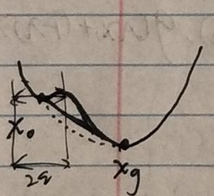
Examine : $\xi^TA^TA\xi\geq 0$

for $\forall\xi$, since $\|\xi^TA\|_2^2=\xi^TA^TA\xi\geq 0$

Thus A^TA is semi-definite positive.

Then Hessian of $f(x)$ is positive semi-definite

c)



① Say, there is ~~another~~ a global optimal x_g, for every x in Domain X $f(x_g)\le f(x)$

$f(x_g)\le f(x_0)$

According to the definition of convex function,

$(1-\lambda)f(x_0)+\lambda f(x_g)\geq f(\lambda x_g+(1-\lambda)x_0)$ $\quad\lambda\in(0,1)$

$f(x_0)\geq(1-\lambda)f(x_0)+\lambda f(x_g)$

which means we can always find a λ, $\lambda x_g+(1-\lambda)x_0\in(x_0-\varepsilon,x_0+\varepsilon)$

(Because, it is a line segment between x_0, x_g)

s.t. $f(x_0)\geq f(\lambda x_g+(1-\lambda)x_0)$ contradict to x_0 is a local minimal

which means x_0 is a global minimal.

② if $f(x_g)=f(x_0)$, then $f(x_0)\geq f(\lambda x_g+(1-\lambda)x_0)$

but $f(x_0)\le f(\lambda x_g+(1-\lambda)x_0)$

$f(x_0)=f(\lambda x_g+(1-\lambda)x_0)$ which means x_0 is a global minimal

2.3 C 10 / 10

✓ + 10 pts Correct (proof by contradiction or other)

+ 8 pts Minor logical error

+ 5 pts Reasonable approach, but incomplete

+ 2 pts Some relevant work

+ 0 pts Incorrect/Blank

d) (i) $h(x) = g(x) + f(x)$ is convex

x_1, x_2

~~h(x)~~ ~~h(x_1~~ $\lambda h(x_1) + (1-\lambda) h(x_2) \overset{?}{\geq} h(\lambda x_1 + (1-\lambda) x_2)$

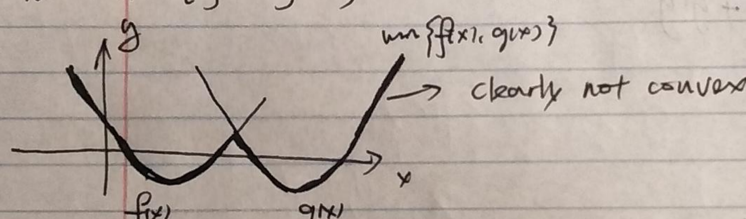
~~Then~~ $\lambda (g(x_1) + f(x_1)) + (1-\lambda)(g(x_2) + f(x_2)) \overset{?}{\geq} g(\lambda x_1 + (1-\lambda) x_2) + h(\lambda x_1 + (1-\lambda) x_2)$

~~Ag(x)~~ which is true because, $g(x)$, $f(x)$ are convex.

$\lambda g(x_1) + (1-\lambda) g(x_2) \geq g(\lambda x_1 + (1-\lambda) x_2)$

$\lambda h(x_1) + (1-\lambda) h(x_2) \geq h(\lambda x_1 + (1-\lambda) x_2)$

(ii) $h(x) = \min\{f(x), g(x)\}$ is not convex



(iii) $h(x) = \max\{f(x), g(x)\}$ is convex

$\lambda \in (0,1)$

$\lambda h(x_1) + (1-\lambda) h(x_2)$ $\qquad$ $h(\lambda x_1 + (1-\lambda) x_2)$

~~max{λf(x_1)+(1-λ)f(x_2), λg(x_1)+(1-λ)g(x_2)}~~ $\qquad$ $\max\{f(\lambda x_1 + (1-\lambda) x_2), g(\lambda x_1 + (1-\lambda) x_2)\}$

$\lambda \max\{f(x_1), g(x_1)\} + (1-\lambda) \max\{f(x_2), g(x_2)\} \geq$

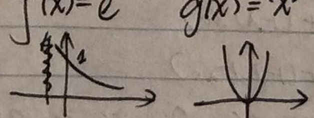
RHS must smaller than LHS

Because $\boxed{\text{if}}$ $\max\{f(\lambda x_1 + (1-\lambda) x_2), g(\lambda x_1 + (1-\lambda) x_2)\} = f(\lambda x_1 + (1-\lambda) x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2)$

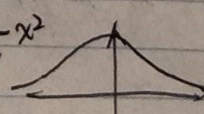
$\leq \lambda \max\{f(x_1), g(x_1)\} + (1-\lambda) \max\{f(x_2), g(x_2)\}$

(iv) $h(x) = f(g(x))$ is not convex

Say $f(x) = e^{-x}$ $\quad g(x) = x^2$



But e^{-x^2} not convex



(looks like Gaussian)

2.4 d 10 / 10

✓ + 2.5 pts part 1 correct

✓ + 2.5 pts part 2 correct

✓ + 2.5 pts part 3 correct

✓ + 2.5 pts part 4 correct

+ 0 pts no part correct

HO5-3

a) $A = U\Sigma V^T$

$$= \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_d \\ & 0 & & \end{bmatrix} \begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_d^T \end{bmatrix}$$

$$= \left(\begin{bmatrix} u_1 & 0 \end{bmatrix} + \begin{bmatrix} \vec{0} & \vec{u}_2 & 0 \end{bmatrix} + \cdots + \begin{bmatrix} 0 & \vec{u}_n \end{bmatrix} \right) \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \ddots \ \sigma_d \\ & 0 & \end{bmatrix} \left(\begin{bmatrix} v_1^T \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \vec{v}_2^T \\ 0 \end{bmatrix} + \cdots + \begin{bmatrix} 0 \\ \vec{v}_n \end{bmatrix} \right)$$

$$\begin{bmatrix} 0 & u_i & 0 \end{bmatrix} \begin{bmatrix} 0 \\ v_i^T \\ 0 \end{bmatrix} = u_i \cdot v_i^T \qquad \begin{bmatrix} 0 & u_i & 0 \end{bmatrix} \begin{bmatrix} 0 \\ v_j^T \\ 0 \end{bmatrix} = 0$$

$i \neq j$

Thus $A = \sum_{i=1}^{d} \sigma_i u_i v_i^T$

b) $A = U\Sigma V^T \quad A^TA = V\Sigma^2V^T \quad \Sigma^2 = \Sigma_{d\times n}\Sigma_{n\times d} = \begin{bmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_d^2 \end{bmatrix}$

$A^TAV = V\Sigma^2$, which means

A^TA has eigenvalue $\lambda_i = \sigma_i^2$ with associated eigenvector v_i

$AA^T = U\Sigma^2U^T \quad A^TAU = U\Sigma^2 \quad \Sigma^2 = \Sigma_{n\times d}\Sigma_{d\times n} = \begin{bmatrix} \sigma_1^2 & & & \\ & \ddots & & \\ & & \sigma_d^2 & \\ & & & 0 \end{bmatrix} \} n\times d$

AA^T has eigenvalue $\lambda_i = \sigma_i^2$ with associated eigenvector u_i

c) $\sigma_1(A) = \max_{\substack{u:\|u\|_2=1 \\ v:\|v\|_2=1}} u^TAv \qquad A = \sum_{i=1}^{d} \sigma_i u_i v_i^T$

(Suppose $\sigma_1 \geq \sigma_2 \cdots \geq \sigma_d$)

Assume $u^* = \sum_{i=1}^{n} a_i u_i \quad v^* = \sum_{i=1}^{d} b_i v_i$

$\sigma_1(A) = \max \left(\sum_{i=1}^{n} a_i u_i\right) A \left(\sum_{i=1}^{d} b_i v_i\right)$ consider, $\sum_{i=1}^{d} a_i b_i \sigma_i \leq \|\vec{t}\|_1 \|\vec{\sigma}\|_\infty$

$= \max \sum_{i=1}^{d} a_i b_i \sigma_i$ where $t_i = a_i b_i$, $\vec{\sigma} = [\sigma_1 \sigma_2 \cdots \sigma_d]^T$

$\vec{a} = (|a_1|, |a_2|, \cdots, |a_d|)^T \quad \vec{b} = (|b_1|, |b_2|, \cdots |b_d|)^T$

$\sum_{i=1}^{d} a_i b_i \sigma_i \leq \|\vec{t}\|_1 \|\vec{\sigma}\|_\infty = \sum_{i=1}^{d} |t_i| \sigma_1 = |a^T b| \sigma_1 \leq \|a\|_2 \|b\|_2 \sigma_1$

$\because \|\vec{u}^*\|_2^2 = \|\sum_{i=1}^{n} a_i u_i\|_2^2 = \left\| \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \right\|_2^2 = \left\| \begin{bmatrix} |a_1| \\ |a_2| \\ \vdots \\ |a_n| \end{bmatrix} \right\|_2^2 = \|\vec{a}\|_2^2 = 1$

3.1 a **10 / 10**

✓ + **10 pts** Correct

+ **0 pts** Insufficient justification - Did not use SVD, made unjustified logical jumps

HOS-3

a) $A = U\Sigma V^T$

$$= \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_d \\ & 0 & & \end{bmatrix} \begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_d^T \end{bmatrix}$$

$$= \left(\begin{bmatrix} u_1 & 0 \end{bmatrix} + \begin{bmatrix} \vec{0} & \vec{u}_2 & 0 \end{bmatrix} + \cdots + \begin{bmatrix} 0 & \vec{u}_n \end{bmatrix} \right) \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \ddots \ \sigma_d \\ & 0 & \end{bmatrix} \left(\begin{bmatrix} v_1^T \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \vec{v}_2^T \\ 0 \end{bmatrix} + \cdots + \begin{bmatrix} 0 \\ \vec{v}_n \end{bmatrix} \right)$$

$$\begin{bmatrix} 0 & u_i & 0 \end{bmatrix} \begin{bmatrix} 0 \\ v_i^T \\ 0 \end{bmatrix} = u_i \cdot v_i^T \qquad \begin{bmatrix} 0 & u_i & 0 \end{bmatrix} \begin{bmatrix} 0 \\ v_j^T \\ 0 \end{bmatrix}_{i \neq j} = 0$$

Thus $A = \sum_{i=1}^{d} \sigma_i u_i v_i^T$

b) $A = U\Sigma V^T \quad A^TA = V\Sigma^2 V^T \quad \Sigma^2 = \Sigma_{d \times n} \Sigma_{n \times d} = \begin{bmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_d^2 \end{bmatrix}$

$A^TAV = V\Sigma^2$, which means

A^TA has eigenvalue $\lambda_i = \sigma_i^2$ with associated eigenvector v_i

$AA^T = U\Sigma^2 U^T \quad A^TAU = U\Sigma^2 \quad \Sigma^2 = \Sigma_{n \times d} \Sigma_{d \times n} = \begin{bmatrix} \sigma_1^2 & & & \\ & \ddots & & \\ & & \sigma_d^2 & \\ & & & 0 \end{bmatrix} \}_{n-d}$

AA^T has eigenvalue $\lambda_i = \sigma_i^2$ with associated eigenvector u_i

c) $\sigma_1(A) = \max\limits_{\substack{u:\|u\|_2=1 \\ v:\|v\|_2=1}} u^TAv \qquad A = \sum_{i=1}^{d} \sigma_i u_i v_i^T$

(suppose $\sigma_1 \geq \sigma_2 \cdots \geq \sigma_d$)

Assume $u^* = \sum_{i=1}^{n} a_i u_i \quad v^* = \sum_{i=1}^{d} b_i v_i$

$\sigma_1(A) = \max \left(\sum_{i=1}^{n} a_i u_i\right) A \left(\sum_{i=1}^{d} b_i v_i\right)$ consider, $\sum_{i=1}^{d} a_i b_i \sigma_i \leq \|\vec{t}\|_1 \|\vec{\sigma}\|_\infty$

$= \max \sum_{i=1}^{d} a_i b_i \sigma_i$ where $t_i = a_i b_i$, $\vec{\sigma} = [\sigma_1 \sigma_2 \cdots \sigma_d]^T$

$\therefore \vec{a} = (|a_1|, |a_2|, \cdots, |a_d|)^T \quad \vec{b} = (|b_1|, |b_2|, \cdots |b_d|)^T$

$\sum_{i=1}^{d} a_i b_i \sigma_i \leq \|\vec{t}\|_1 \|\vec{\sigma}\|_\infty = \sum_{i=1}^{d} |t_i| \sigma_1 = |\vec{a}^T b| \sigma_1 \leq \|a\|_2 \|b\|_2 \sigma_1$

$\because \|\vec{u}^*\|_2^2 = \|\sum_{i=1}^{n} a_i u_i\|_2^2 = \left\| \begin{bmatrix} u_1 u_2 \cdots u_n \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \right\|_2^2 = \left\| \begin{bmatrix} |a_1| \\ |a_2| \\ \vdots \\ |a_n| \end{bmatrix} \right\|_2 = \|\vec{a}\|_2^2 = 1$

3.2 b **10 / 10**

✓ + **10 pts** correct

+ **8 pts** almost correct

+ **5 pts** halfway

+ **2 pts** minimal effort

+ **0 pts** no effort

HO5-3

a) $A = U\Sigma V^T$

$$= \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_d \\ & 0 & & \end{bmatrix} \begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_d^T \end{bmatrix}$$

$$= \left(\begin{bmatrix} u_1 & 0 \end{bmatrix} + \begin{bmatrix} \vec{0} & \vec{u_2} & 0 \end{bmatrix} + \cdots + \begin{bmatrix} 0 & \vec{u_n} \end{bmatrix} \right) \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \sigma_d \\ & 0 & & \end{bmatrix} \left(\begin{bmatrix} v_1^T \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \vec{v_2}^T \\ 0 \end{bmatrix} + \cdots + \begin{bmatrix} 0 \\ \vec{v_n} \end{bmatrix} \right)$$

$$\begin{bmatrix} 0 & u_i & 0 \end{bmatrix} \begin{bmatrix} 0 \\ v_i^T \\ 0 \end{bmatrix} = u_i \cdot v_i^T \qquad \begin{bmatrix} 0 & u_i & 0 \end{bmatrix} \begin{bmatrix} 0 \\ v_j^T \\ 0 \end{bmatrix} = 0 \quad i \neq j$$

Thus $A = \sum_{i=1}^{d} \sigma_i u_i v_i^T$

b) $A = U\Sigma V^T$ $\quad A^TA = V\Sigma^2 V^T \quad \Sigma^2 = \Sigma_{d\times n}\Sigma_{n\times d} = \begin{bmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_d^2 \end{bmatrix}$

$A^TAV = V\Sigma^2$, which means

A^TA has eigenvalue $\lambda_i = \sigma_i^2$ with associated eigenvector v_i

$AA^T = U\Sigma^2 U^T \quad A^TAU = U\Sigma^2 \quad \Sigma^2 = \Sigma_{n\times d}\Sigma_{d\times n} = \begin{bmatrix} \sigma_1^2 & & & \\ & \ddots & & \\ & & \sigma_d^2 & \\ & & & 0 \end{bmatrix} \Big\} n\times d$

AA^T has eigenvalue $\lambda_i = \sigma_i^2$ with associated eigenvector u_i

c) $\sigma_1(A) = \max_{\substack{u: \|u\|_2 = 1 \\ v: \|v\|_2 = 1}} u^T A v \qquad A = \sum_{i=1}^{d} \sigma_i u_i v_i^T$

(suppose $\sigma_1 \geq \sigma_2 \cdots \geq \sigma_d$)

Assume $u^* = \sum_{i=1}^{n} a_i u_i \quad v^* = \sum_{i=1}^{d} b_i v_i$

$\sigma_1(A) = \max \left(\sum_{i=1}^{n} a_i u_i\right) A \left(\sum_{i=1}^{d} b_i v_i\right)$ $\quad$ consider, $\sum_{i=1}^{d} a_i b_i \sigma_i \leq \|\vec{t}\|_1 \|\vec{\sigma}\|_\infty$

$= \max \sum_{i=1}^{d} a_i b_i \sigma_i$ $\quad$ where $t_i = a_i b_i$, $\vec{\sigma} = [\sigma_1 \sigma_2 \cdots \sigma_d]^T$

$\vec{a} = (|a_1|, |a_2|, \cdots, |a_d|)^T \quad \vec{b} = (|b_1|, |b_2|, \cdots |b_d|)^T$

$\sum_{i=1}^{d} a_i b_i \sigma_i \leq \|\vec{t}\|_1 \|\vec{\sigma}\|_\infty = \sum_{i=1}^{d} |t_i| \sigma_1 = |\vec{a}^T \vec{b}| \sigma_1 \leq \|a\|_2 \|b\|_2 \sigma_1$

$\because \|\vec{u}^*\|_2^2 = \|\sum_{i=1}^{n} a_i u_i\|_2^2 = \left\| \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \right\|_2^2 = \left\| \begin{bmatrix} |a_1| \\ |a_2| \\ \vdots \\ |a_n| \end{bmatrix} \right\|_2^2 = \|\vec{a}\|_2^2 = 1$

Also $\|\vec{u}\|_2 = \|\vec{b}\|_2 = 1$ $\qquad \vec{t} = \begin{bmatrix} t_1 \\ \vdots \\ t_d \end{bmatrix}$, $t_i = a_i b_i$

Thus $\sum_{i=1}^{d} a_i b_i \sigma_i \le \|\vec{t}\|_1 \|\vec{\sigma}\|_\infty \le \|\vec{a}\|_2 \|\vec{b}\|_2 \|\vec{\sigma}\|_\infty$

Now, we want a maximal $\sum_{i=1}^{d} a_i b_i \sigma_i$, when the 2 inequality holds.

The second holds when $\vec{a} = \alpha \vec{b}$ $\quad \alpha$ is a scalar.

First holds when $\|\vec{\sigma}\|_\infty = |\sigma_1|$ $(i=1)$

Set $t_1 = \|t\|_1$ $\quad t_j = 0$, $j \ne 1$

Because $\|\vec{a}\|_2 = \|\vec{b}\|_2 = 1$ $\quad t_1 = |a_1||b_1|$

Then, $|a_1| = |b_1| = 1$ $\quad a_j = b_j = 0$, $j \ne 1$

We have, when $\vec{a} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ $\vec{b} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, $\max \sum_{i=1}^{d} a_i b_i \sigma_i = \|\vec{t}\|_1 \|\vec{\sigma}\|_\infty = \|\vec{a}\|_2 \|\vec{b}\|_2 \sigma_1 = \sigma_1$ clear.

$\vec{u}^* = U \cdot \vec{a} = \vec{u}_1 \qquad \vec{v}^* = V \cdot \vec{b} = \vec{v}_1$

Finally, we prove, $\sigma_1(A)$ is the maximum singular value of A.

If A has a unique maximum, (u^*, v^*) are first left/right singular vector.

d) $$\rho = \max_{a,b} \rho(a^T X, b^T Y) = \max_{a,b} \frac{E[a^T X Y^T b]}{\sqrt{E[a^T X X^T a]\, E[b^T Y Y^T b]}}$$

$1\times n$ $n\times d$ $d\times 1$: $E[a^T A b] = a^T E[A] b$

$$= \max_{a,b} \frac{a^T E[XY^T] b}{\sqrt{a^T E[XX^T] a}\sqrt{b^T E[YY^T] b}}$$

$$= \max_{a,b} \frac{a^T \Sigma_{XY} b}{(a^T \Sigma_{XX} a)^{1/2} (b^T \Sigma_{YY} b)^{1/2}}$$

If $$\rho^* = \rho(a^{*T} X, b^{*T} Y) = \left.\frac{a^{*T} \Sigma_{XY} b^*}{(a^{*T} \Sigma_{XX} a^*)^{1/2} (b^{*T} \Sigma_{YY} b^*)^{1/2}}\right|_{a=a^*,\, b=b^*}$$

$$\rho(\alpha a^{*T} X, \beta b^{*T} Y) = \frac{\beta \alpha a^{*T} \Sigma_{XY} b^*}{(\alpha^2 a^{*T} \Sigma_{XX} a^*)^{1/2} (\beta^2 b^{*T} \Sigma_{YY} b^*)^{1/2}} = \frac{a^{*T} \Sigma_{XY} b^*}{(a^{*T} \Sigma_{XX} a^*)^{1/2} (b^{*T} \Sigma_{YY} b^*)^{1/2}} = \rho^*$$

Thus $(\alpha a^*, \beta b^*)$ is a maximizer for any $\alpha, \beta > 0$.

3.3 C 10 / 10

✓ + **10 pts** correct (doesn't have to use hint2)

+ **9 pts** almost correct
+ **7 pts** realize the application of hint1
+ **5 pts** realize the application of vector
+ **3 pts** say things about norm vector
+ **0 pts** incorrect

$a_i = \alpha b_i$ $|a_1|b_1| = |a_{b1}|$

Also $\|\vec{u}\|_2 = \|\vec{b}\|_2 = 1$ $\qquad \vec{t} = \begin{bmatrix} t_1 \\ \vdots \\ t_d \end{bmatrix}$, $t_i = a_i b_i$

Thus $\sum_{i=1}^{d} a_i b_i \sigma_i \le \|\vec{t}\|_1 \|\vec{\sigma}\|_\infty \le \|\vec{a}\|_2 \|\vec{b}\|_2 \|\vec{\sigma}\|_\infty$

Now, we want a maximal $\sum_{i=1}^{d} a_i b_i \sigma_i$, when the 2 inequality holds.

The second holds when $\vec{a} = \alpha \vec{b}$ α is a scalar.

First holds when $\|\vec{\sigma}\|_\infty = |\sigma_1|$ $(i=1)$

Set $t_1 = \|t\|_1$ $t_j = 0$, $j \ne 1$

Because $\|\vec{a}\|_2 = \|\vec{b}\|_2 = 1$ $t_1 = |a_1||b_1|$

Then, $|a_1| = |b_1| = 1$ $a_j = b_j = 0$, $j \ne 1$

We have, when $\vec{a} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ $\vec{b} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, $\max \sum_{i=1}^{d} a_i b_i \sigma_i = \|\vec{t}\|_1 \|\vec{\sigma}\|_\infty = \|\vec{a}\|_2 \|\vec{b}\|_2 \sigma_1 = \underset{\text{clear}}{\sigma_1}$.

$\vec{u}^* = U \cdot \vec{a} = \vec{u}_1$ $\qquad \vec{v}^* = V \cdot \vec{b} = \vec{v}_1$

Finally, we prove, $\sigma_1(A)$ is the maximum singular value of A.

If A has a unique maximum, (u^*, v^*) are first left / right singular vector.

d) $\rho = \max_{a,b} \rho(a^T X, b^T Y) = \max_{a,b} \frac{E[a^T X Y^T b]}{\sqrt{E[a^T X X^T a]\, E[b^T Y Y^T b]}}$ $\qquad \underset{1\times n \;\; n\times d \;\; d\times 1}{E[a^T A b]} = a^T E[A] b$

$$= \max_{a,b} \frac{a^T E[X Y^T] b}{\sqrt{a^T E[X X^T] a}\sqrt{b^T E[Y Y^T] b}}$$

$$= \max_{a,b} \frac{a^T \Sigma_{XY} b}{(a^T \Sigma_{XX} a)^{1/2} (b^T \Sigma_{YY} b)^{1/2}}.$$

If $\rho^* = \rho(a^{*T} X, b^{*T} Y) = \left.\frac{a^{*T} \Sigma_{XY} b^*}{(a^{*T} \Sigma_{XX} a^*)^{1/2} (b^{*T} \Sigma_{YY} b^*)^{1/2}}\right|_{a=a^*, b=b^*}$

$$\rho(\alpha a^{*T} X, \beta b^{*T} Y) = \frac{\beta \alpha a^{*T} \Sigma_{XY} b^*}{(\alpha^2 a^{*T} \Sigma_{XX} a^*)^{1/2} (\beta^2 b^{*T} \Sigma_{YY} b^*)^{1/2}} = \frac{a^{*T} \Sigma_{XY} b^*}{(a^{*T} \Sigma_{XX} a^*)^{1/2} (b^{*T} \Sigma_{YY} b^*)^{1/2}} = \rho^*$$

Thus $(\alpha a^*, \beta b^*)$ is a maximizer for any $\alpha, \beta > 0$.

3.4 d 10 / 10

✓ + 10 pts Correct/Partially Correct

+ 0 pts Incorrect

HoG-3

e) i) $\rho = \max_{a,b\in\mathbb{R}^d} \frac{a^T\Sigma_{XY}b}{(a^T\Sigma_{XX}a)^{1/2}(b^T\Sigma_{YY}b)^{1/2}}$

$\begin{cases} \Sigma_{XX} = U_X\Sigma_X U_X^T \\ \Sigma_{YY} = U_Y\Sigma_Y U_Y^T \end{cases}$ $a^T\Sigma_{XX}a = a^TU_X\Sigma_X^{1/2}U_X^T(U_X\Sigma_X^{\frac{1}{2}}U_X^Ta)$ Set $a' = U_X\Sigma_X^{\frac{1}{2}}U_X^Ta = \Sigma_{XX}^{\frac{1}{2}}a$, $a = \Sigma_{XX}^{-\frac{1}{2}}a'$

Same trick, $b' = \Sigma_{YY}^{\frac{1}{2}}b$, $b = \Sigma_{YY}^{-\frac{1}{2}}b'$

Mind Σ_X, Σ_{XX}

$\Downarrow$

diagonal, all element ≥ 0

$\Sigma_X^{\frac{1}{2}} = \begin{bmatrix} \sigma_1^{\frac{1}{2}} & & \\ & \ddots & \\ & & \sigma_d^{\frac{1}{2}} \end{bmatrix}$ $\rho = \max_{a,b\in\mathbb{R}^d} \frac{a'^T\Sigma_{XX}^{-\frac{1}{2}}\cdot\Sigma_{XY}\Sigma_{YY}^{-\frac{1}{2}}b'}{\|a'\|_2 \; \|b'\|_2}$

In part c) we proved that the maximizer is $(\alpha a^*, \beta b^*)$, which means scale doesn't change the optimal of ρ. The above problem can be rewritten as,

$$f(a', b') = \max_{a',b'\in\mathbb{R}^d} a'^T\Sigma_{XX}^{-\frac{1}{2}}\Sigma_{XY}\Sigma_{YY}^{-\frac{1}{2}}b'$$

$$\text{s.t. } \|a'\|_2 = \|b'\|_2 = 1$$

In part c) we also showed $\sigma_1(A) = \max_{\substack{u\,\|u\|_2=1 \\ v\,\|v\|_2=1}} u^TAv$

Then ρ is the singular value of $\Sigma_{XX}^{-\frac{1}{2}}\Sigma_{XY}\Sigma_{YY}^{-\frac{1}{2}}$

ρ^2 must be the eigen value of $\Sigma_{XX}^{-\frac{1}{2}}\Sigma_{XY}\Sigma_{YY}^{-\frac{1}{2}}\Sigma_{YY}^{-\frac{1}{2}}\Sigma_{XY}^T\Sigma_{XX}^{-\frac{1}{2}}$

$\|$

ρ^2 is the maximum eigenvalue of: $\Sigma_{XX}^{-1/2}\Sigma_{XY}\Sigma_Y^{-1}\Sigma_{XY}^T\Sigma_{XX}^{-1/2}$

ii)

In part c) we know $u^* = u_1 \Rightarrow a'^*$ is the left $\overset{\text{largest}}{\vee}$ singular vector of $\Sigma_{XX}^{-\frac{1}{2}}\Sigma_{XY}\Sigma_{YY}^{-\frac{1}{2}}$

a'^* is the eigenvector of $\Sigma_{XX}^{-\frac{1}{2}}\Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{XY}^T\Sigma_{XX}^{-\frac{1}{2}}$

$a^* = \Sigma_{XX}^{-\frac{1}{2}}a'^*$ $\Sigma_{XX}^{\frac{1}{2}}a^*$ is the maximal eigenvector of $\Sigma_{XX}^{-\frac{1}{2}}\Sigma_{XY}\Sigma_{YY}^{-1}$

iii) for the same reason, $\Sigma_{YY}^{\frac{1}{2}}b^*$ is the maximal eigenvector of $\Sigma_{YY}^{-\frac{1}{2}}\Sigma_{XY}^T\Sigma_X^{-\frac{1}{2}}\Sigma_X^{-\frac{1}{2}}\Sigma_{XY}\Sigma_{YY}^{-\frac{1}{2}}$

$\|$

$\Sigma_{YY}^{-\frac{1}{2}}\Sigma_{XY}^T\Sigma_X^{-1}\Sigma_{XY}\Sigma_{YY}^{-\frac{1}{2}}$

3.5 e 10 / 10

- ✓ + **10 pts** Correct
- + **8 pts** Almost Correct
- + **6 pts** Significant Progress
- + **4 pts** Some Progress
- + **2 pts** Trivial Progress
- + **0 pts** Blank or Incorrect

f) If $\mathrm{cov}(X_i, Y_j) = 0 = \dfrac{E[XY^T]}{\sqrt{E[XX^T]E[YY^T]}} = 0$

Then, $\rho = \max_{a,b} \rho(a^T X, b^T Y)$ $E[XY^T]$ is 0 matrix

$$= \frac{a^T E[XY^T] b}{(a^T E[XX^T] a)^{\frac{1}{2}} (b^T E[YY^T] b)^{\frac{1}{2}}}$$

No matter what value a, b take ρ will be zero. "vanilla" CCA becomes useless.

set $Y^2 = Z$ solve

$$\rho = \max_{a,b \in R} \rho(a^T X, b^T Z)$$

① get Σ_{XX} Σ_{ZZ} Σ_{XZ}

② using conclusion from part e, get ρ, a^* b^*

3.6 f 10 / 10

✓ + **10 pts** **Correct argument about uselessness of CCA and correct modification of CCA**

+ **8 pts** Minor error in one part

+ **5 pts** Only one part correct (or both partially correct)

+ **2 pts** Only one part partially correct

+ **0 pts** Incorrect or blank

H05-4

a) $\Sigma_{XX} = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})(x_i - \bar{x})^T$ where $\bar{x} = \sum_{i=1}^{n} x_i$ $\quad x_i = \begin{bmatrix} x_1^{(i)} \\ \vdots \\ x_d^{(i)} \end{bmatrix}$

$$\Sigma_{YY} = \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y})(y_i - \bar{y})^T$$

$$\Sigma_{XY} = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})^T$$

b) code

c) code

d) $\min_{w} \|(\underbrace{X}_{956\times675}\underbrace{P_k}_{675\times k})\underbrace{w}_{k\times675} - \underbrace{Y}_{956\times675}\|_2^2 + \lambda\|w\|_2^2$

$$w = \left((XP_k)^T XP_k + \lambda I\right)^{-1}(XP_k)^T Y$$

e) code

$E[XX^T]$

4.1 a **3 / 10**

- ✓ + **3 pts** Explain how to estimate covariance matrices
- \+ **3 pts** Covariance Matrices in Code
- \+ **2 pts** Standardized data AND subtracted mean
- \+ **2 pts** Scaled at the end
- \+ **0 pts** Wrong/Missing

HO5-4

a) $\Sigma_{XX} = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})(x_i - \bar{x})^T$ where $\bar{x} = \sum_{i=1}^{n} x_i$ $\quad x_i = \begin{bmatrix} x_1^{(i)} \\ \vdots \\ x_d^{(i)} \end{bmatrix}$

$$\Sigma_{YY} = \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y})(y_i - \bar{y})^T$$

$$\Sigma_{XY} = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})^T$$

b) code

c) code

d) $\min_{w} \|(\underset{956\times675}{X}\underset{675\times k}{P_k})\underset{k\times675}{w} - \underset{956\times675}{Y}\|_2^2 + \lambda\|w\|_2^2$

$$w = \left((XP_k)^T XP_k + \lambda I\right)^{-1} (XP_k)^T Y$$

e) code

$E[XX^T]$

4.2 b 0 / 10

- ✓ + **0 pts** **Blank/Wrong**
- + **2 pts** Minimal Progress
- + **5 pts** Halfway
- + **8 pts** Minor Error
- + **10 pts** Correct

H05-4

a) $\Sigma_{XX} = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})(x_i - \bar{x})^T$ where $\bar{x} = \sum_{i=1}^{n} x_i$ $\quad x_i = \begin{bmatrix} x_1^{(i)} \\ \vdots \\ x_d^{(i)} \end{bmatrix}$

$$\Sigma_{YY} = \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y})(y_i - \bar{y})^T$$

$$\Sigma_{XY} = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})^T$$

b) code

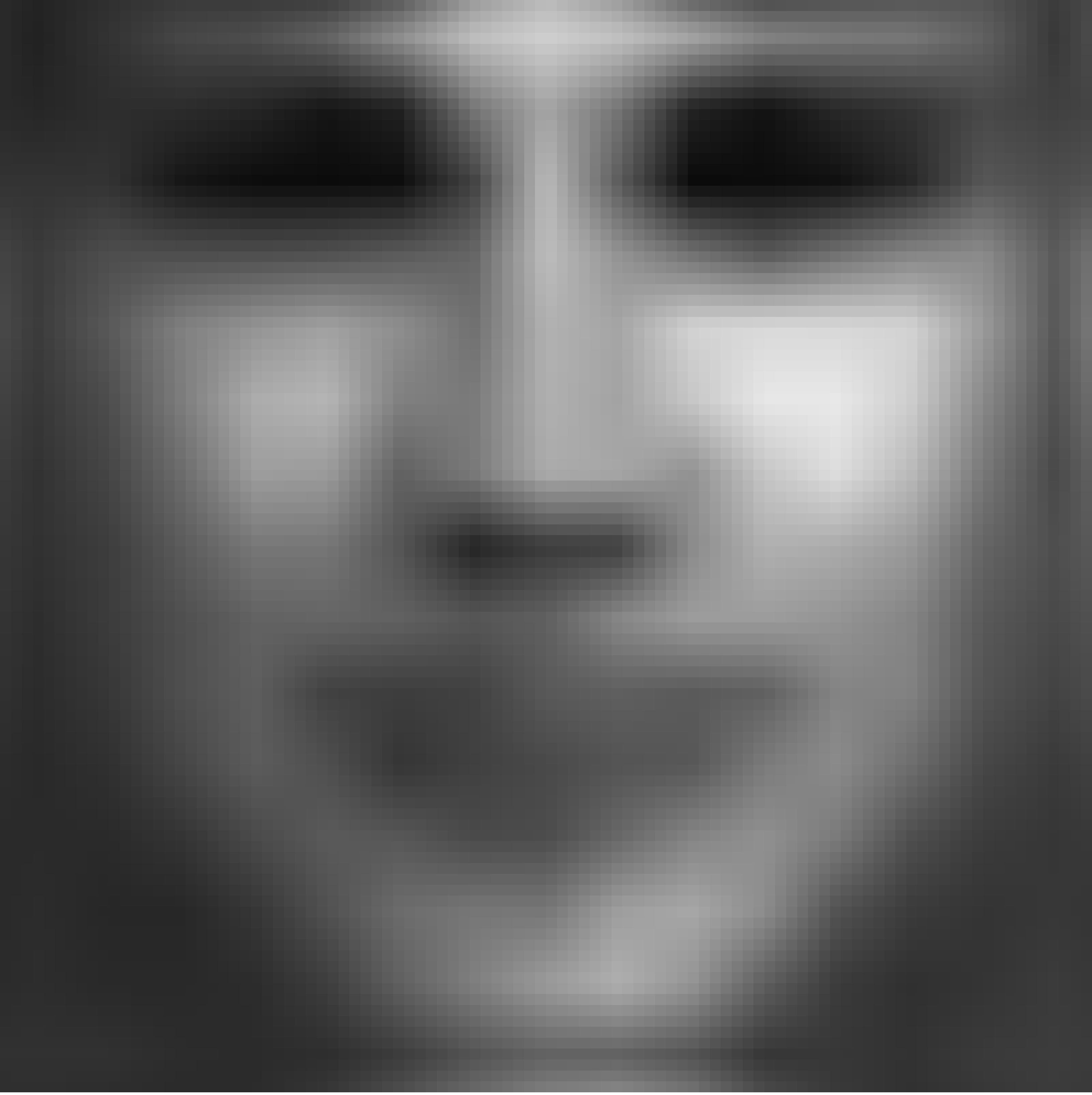
c) code

d) $\min_{\omega} \|(\underbrace{X}_{956\times675} \underbrace{P_k}_{675\times k})\underbrace{\omega}_{k\times675} - \underbrace{Y}_{956\times675}\|_2^2 + \lambda\|\omega\|_2^2$

$$\omega = \left((XP_k)^T XP_k + \lambda I\right)^{-1} (XP_k)^T Y$$

e) code

$E[XX^T]$



4.3 C 10 / 10

✓ + 10 pts Correct or Minor Error

+ 8 pts Image is off

+ 5 pts Partial

+ 2 pts Reasonable attempt

+ 0 pts Incorrect or Missing

H05-4

a) $\Sigma_{XX} = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})(x_i - \bar{x})^T$ where $\bar{x} = \sum_{i=1}^{n} x_i$ $x_i = \begin{bmatrix} x_1^{(i)} \\ \vdots \\ x_d^{(i)} \end{bmatrix}$

$\Sigma_{YY} = \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y})(y_i - \bar{y})^T$

$\Sigma_{XY} = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})^T$

b) code

c) code

d) $\min_{w} \|(\underbrace{X}_{956\times 675}\underbrace{P_k}_{675\times k})\underbrace{w}_{k\times 675} - \underbrace{Y}_{956\times 675}\|_2^2 + \lambda \|w\|_2^2$

$w = \left((XP_k)^T XP_k + \lambda I\right)^{-1} (XP_k)^T Y$

e) code

$E[XX^T]$

4.4 d 0 / 10

+ **10 pts** Correct

+ **5 pts** An attempt, but plot is off

✓ + **0 pts** **No Answer / Incorrect**

H05-4

a) $\Sigma_{XX} = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})(x_i - \bar{x})^T$ where $\bar{x} = \sum_{i=1}^{n} x_i$ $\quad x_i = \begin{bmatrix} x_1^{(i)} \\ \vdots \\ x_d^{(i)} \end{bmatrix}$

$\Sigma_{YY} = \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y})(y_i - \bar{y})^T$

$\Sigma_{XY} = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})^T$

b) code

c) code

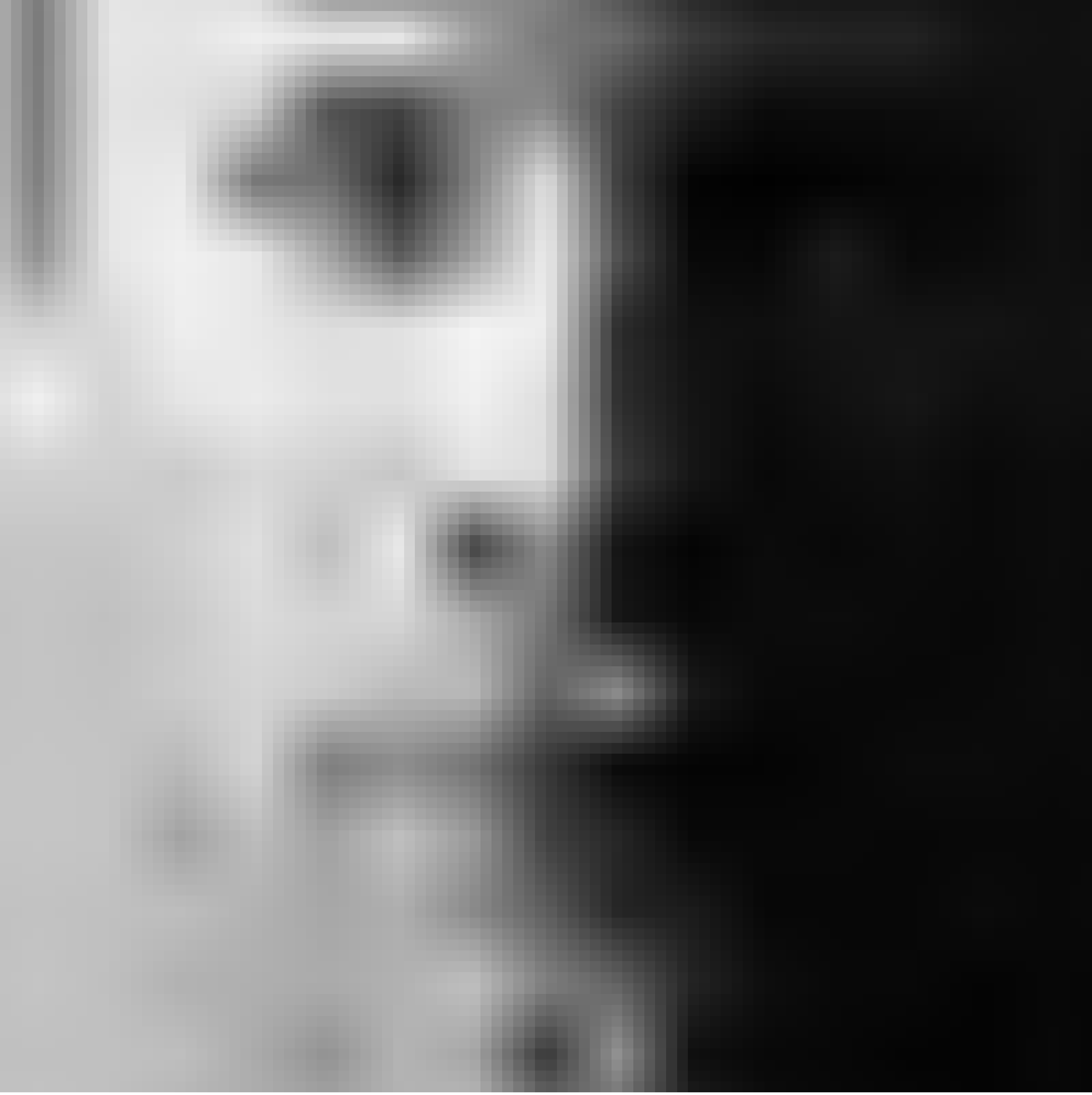
d) $\min_{w} \|(X P_k) w - Y\|_2^2 + \lambda \|w\|_2^2$

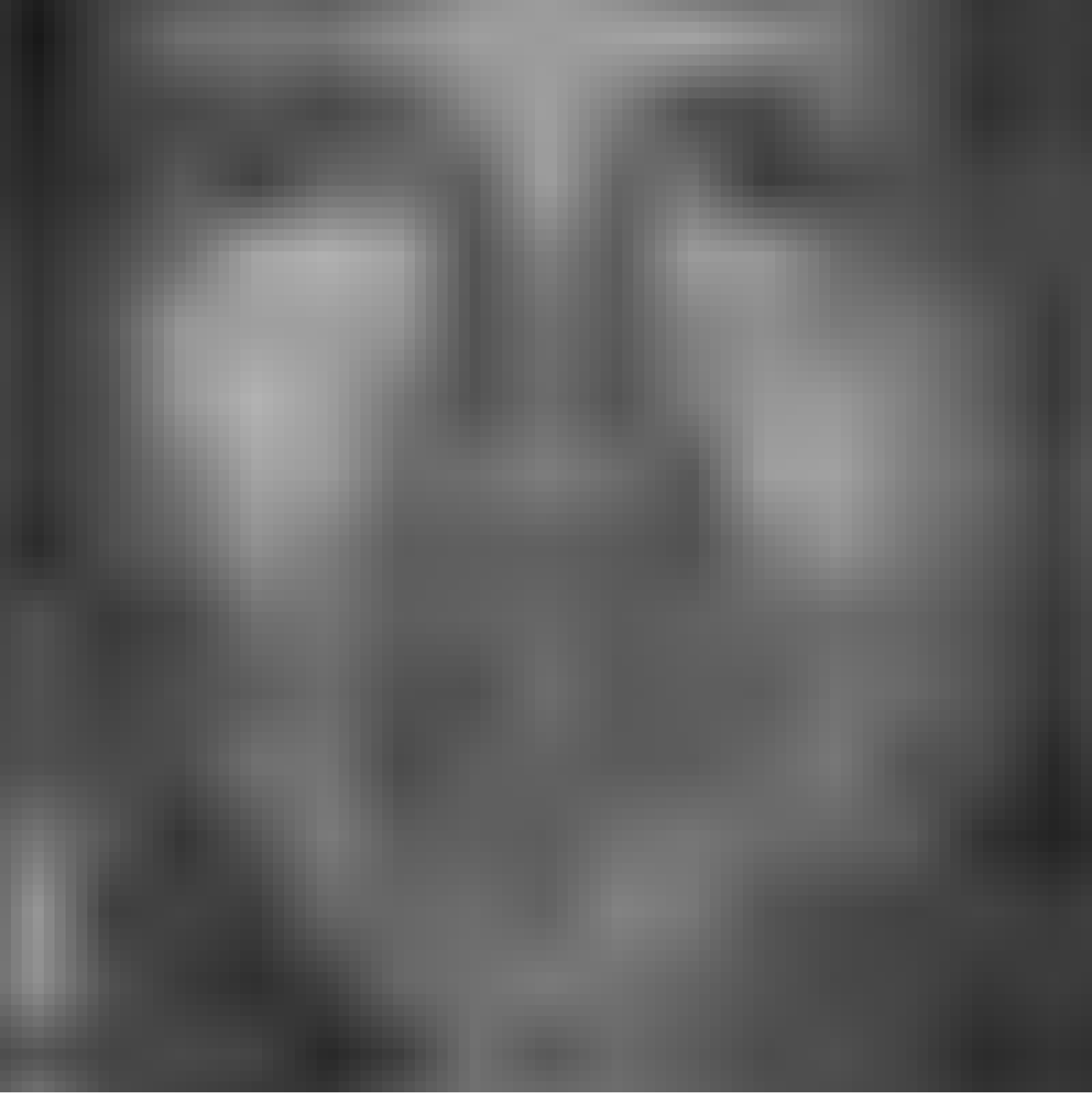
956x675 675xk kx675 956x675

$w = \left((XP_k)^T XP_k + \lambda I\right)^{-1} (XP_k)^T Y$

e) code

$E[XX^T]$





4.5 e 10 / 10

- ✓ + **10 pts** **Complete**
- + **8 pts** Close to complete
- + **5 pts** Partial
- + **2 pts** Reasonable Attempt
- + **0 pts** Wrong or Missing

SweeTango
USA
3603


SweeTango
USA


Chiquita
#4011


#4011
Chiquita
Mexico


EACH
Every Day!

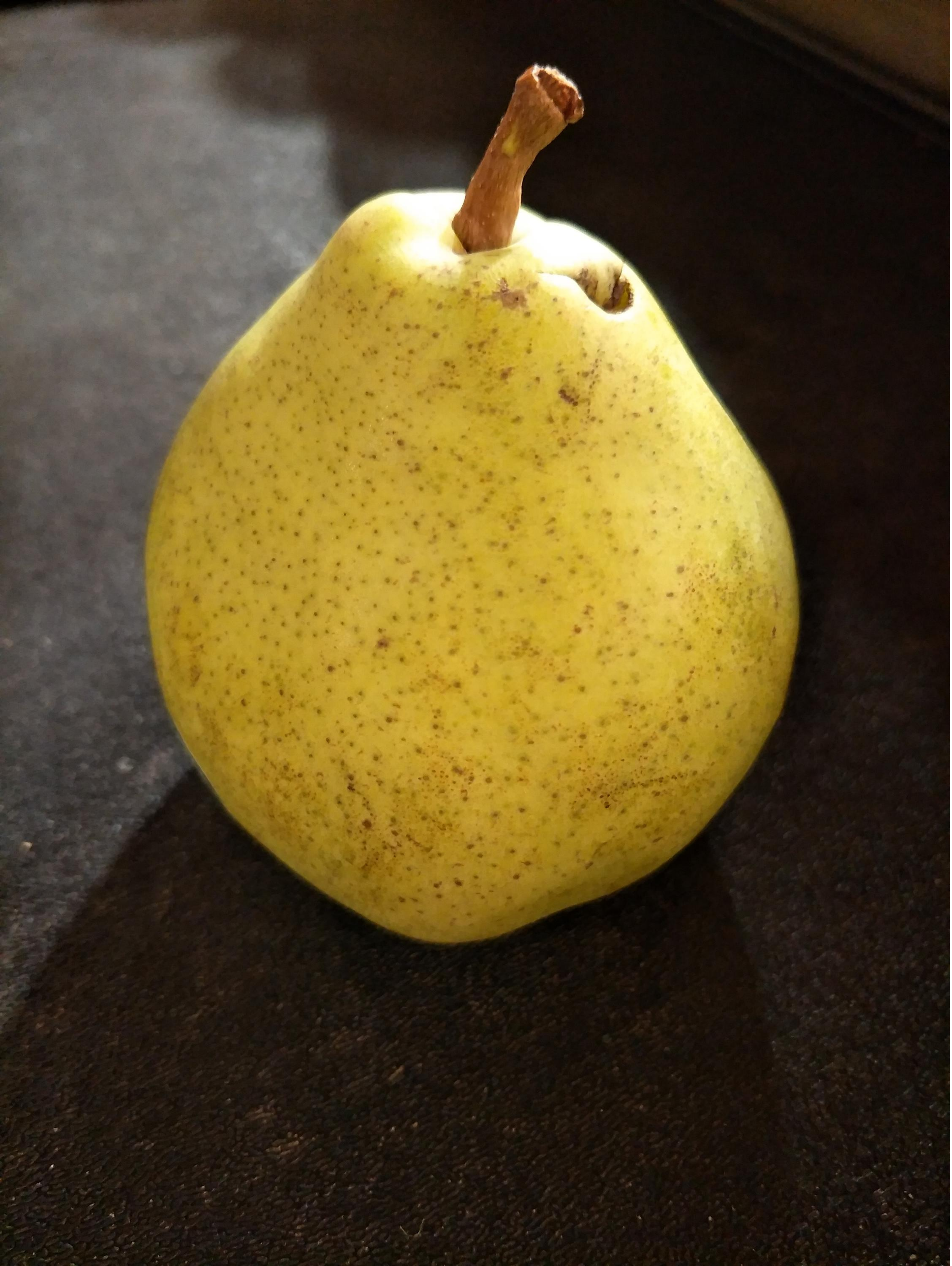

PLU #4333
399
ea
$3.99
ORGANIC MANGOS
EACH
PLU #94051
199
Every Day!
ea
$1.99


4378
CALIFORNIA
CALIFORNIE
U.S.A./É-U.


4378
CALIFORNIA
CALIFORNIE


FRESH


SCULLY
COMICE
4414
USA




CALIFORNIE
USA
4040


ORGANIC

5 Fruits! 0 / 0

✓ + **0 pts** Fruits!

HO5-6

Holder's Inequality : $\|fg\|_1 \leq \|f\|_p \|g\|_q$ for $\frac{1}{p}+\frac{1}{q}=1$

for our problem $\underbrace{|a^T b|}_{\text{number 1-norm is itself}} \leq \|a\|_1 \|b\|_\infty$ $\quad \frac{1}{1}+\frac{1}{\infty}=1$

Proof: Lemma: $ab \leq \frac{a^p}{p}+\frac{b^q}{q}$

$A=\|f\|_p \quad B=\|g\|_q \quad A, B \neq 0.$

$a=\frac{|f(x)|}{A} \quad b=\frac{|g(x)|}{B} \quad ab=\frac{|f(x)g(x)|}{AB} \leq \frac{|f(x)|^p}{pA^p}+\frac{|g(x)|^q}{qB^q}=\frac{a^p}{p}+\frac{b^q}{q}$

$$\frac{1}{AB}\int |f(x)g(x)|\,d\mu \leq \frac{1}{pA^p}\int |f|^p d\mu + \frac{1}{qB^q}\int |g|^q d\mu$$

$A^p=\int |f|^p d\mu \quad B^q=\int |g|^q d\mu$

$$\frac{1}{\|f\|_p\|g\|_q}\|fg\|_1 \leq \frac{1}{p}+\frac{1}{q}=1$$

$$\|fg\|_1 \leq \|f\|_p\|g\|_q$$

references: www.math.cornell.edu/~erin/analysis/lectures.pdf

6 Your Own Question 10 / 10

✓ + **10 pts** **Correct**

+ **5 pts** Question, no solution

+ **0 pts** No question or no effort