Team Evaluation

1. Brendan Rizzo
   a. Worked with me, Eden Seo, to solve problem 2 part a and b, which is exercise 3.3
2. Richard Huang
   a. Solved problem 1 part a by himself and explained to us how he got the answer.
3. Jiashang Cao
   a. Solved problem 1 part b
4. **Eden Seo**
   a. Worked with Brendan to solve problem 2 part a and b, and solved part c and d by myself.

1.

a. $Derive\ the\ optimal\ w\ that\ minimizes\ E_2(w)$

i. $E_2(w) = \|Xw - y\|^2 + \lambda\|w\|^2$

ii. $E_2(w) = ((Xw)' - y')(Xw - y) + \lambda(w' * w) = (X'w' - y')(Xw - y) + \lambda(w' * w)$

iii. $= X'w'Xw - X'w'y - y'Xw + y'y + \lambda(w' * w),\ where\ y'Xw = X'w'y$

iv. $= X'w'Xw - 2X'w'y + y'y + \lambda(w' * w)$

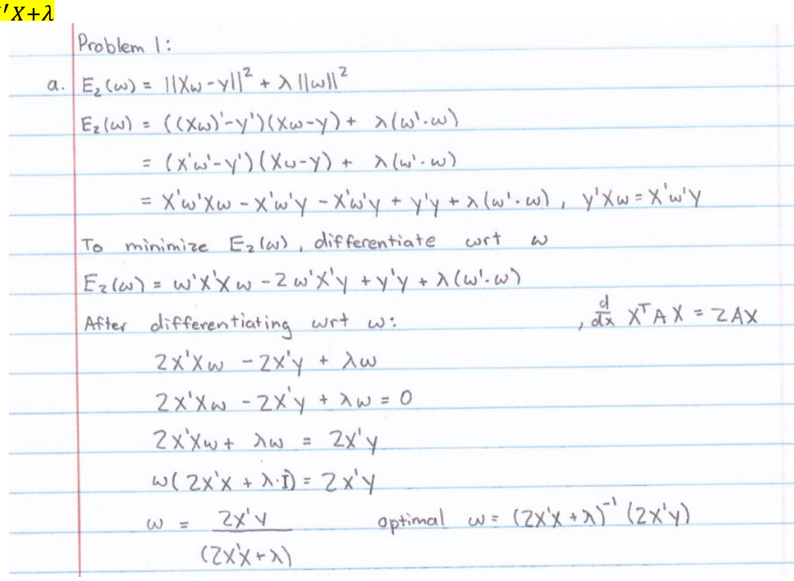
v. $To\ minimize\ E_2(w), differentiate\ with\ respect\ to\ w$

vi. $2X'Xw - 2X'y + \lambda w,\ where \frac{d}{dx}X^T AX = 2AX$

vii. $2X'Xw - 2X'y + \lambda w = 0 \quad => \quad 2X'Xw + \lambda w = 2X'y$

viii. $w(2X'X + \lambda * I) = 2X'y,\ where\ I\ is\ an\ identity\ matrix$

ix. $w = \frac{2X'y}{2X'X+\lambda}$

1. 

Problem 1:

a. $E_2(\omega) = \|X\omega - y\|^2 + \lambda\|\omega\|^2$

$E_2(\omega) = ((X\omega)' - y')(X\omega - y) + \lambda(\omega' \cdot \omega)$

$= (X'\omega' - y')(X\omega - y) + \lambda(\omega' \cdot \omega)$

$= X'\omega'X\omega - X'\omega'y - X'\omega'y + y'y + \lambda(\omega' \cdot \omega),\quad y'X\omega = X'\omega'y$

To minimize $E_2(\omega)$, differentiate wrt ω

$E_2(\omega) = \omega'X'X\omega - 2\omega'X'y + y'y + \lambda(\omega' \cdot \omega)$

After differentiating wrt ω: $\quad , \frac{d}{dx} X^T A X = 2AX$

$2X'X\omega - 2X'y + \lambda\omega$

$2X'X\omega - 2X'y + \lambda\omega = 0$

$2X'X\omega + \lambda\omega = 2X'y$

$\omega(2X'X + \lambda \cdot I) = 2X'y$

$\omega = \frac{2X'y}{(2X'X + \lambda)} \quad$ optimal $\omega = (2X'X + \lambda)^{-1}(2X'y)$

(hand-written work is much easier to see, so here it is.)

b. $From\ 1a, we\ got\ optimal\ \ w = (2X^T X + \lambda)^{-1}(2X^T Y).$

$In\ this\ case, the\ solution\ holds\ only\ when\ 2X^T X + \lambda\ is$

$non\ singular, and\ since\ \lambda > 0\ and\ it\ is\ an\ user\ specific\ parameter,$
$the\ new\ objective\ function\ overcomed\ the\ singularity\ problem$
$of\ X^T X\ using\ the\ parameter\ \ \lambda.$

2. Exercise 3.3 in LFD
   a. Show that H is symmetric.
      i. $H^T = (X(X^T X)^{-1} X^T)^T = (X^T (X^T X)^{-T} X) = (X^T (X X^T)^{-1} X) =$
         $since\ X^T X\ is\ invertible, H^T = (X(X^T X)^{-1} X^T) = H$
   b. Show that $H^K = H\ for\ any\ positive\ integer\ K.$
      i. $H^K = (X(X^T X)^{-1} X^T)^K = \frac{XX^T}{X^T X} * \frac{XX^T}{X^T X} * \dots \frac{XX^T}{X^T X}\ (for\ K\ amount\ of\ time)$
      ii. $since\ X^T X\ is\ invertible, \frac{XX^T}{X^T X}\ is\ equal\ to \frac{X^T X}{X^T X}\ which\ is\ equal\ to\ 1.$
      iii. $So\ no\ matter\ how\ many\ time\ H\ is\ multiplied\ by\ itself, it\ will\ still\ be\ the\ same.$
      iv. $If\ K = 2$
         1. $(X(X^T X)^{-1} X^T) * (X(X^T X)^{-1} X^T) =$
            $X(X^T X)^{-1} (X^T X)(X^T X)^{-1} X^T = X(X^T X)^{-1} X^T$
   c. Show that $Show\ that\ (I - H)^K = I - H\ for\ any\ positive\ integer\ K.$
      i. $if\ K = 2$
         1. $(I - H)^2 = (I - H)(I - H) = II - IH - HI + H^2 = I - 2H +$
            $H^2, and\ since\ H^2 == H, I - 2H + H = I - H$
      ii. $Since\ I^k, where\ K\ is\ a\ positive\ integer, is\ equal\ to\ I\ and\ H^K\ is\ always\ equal\ to\ H,$
      iii. $it\ would\ always\ become\ I - H$
   d. Show that trace(H) = d + 1
      i. $Lets\ say\ that\ A = X(X^T X)^{-1}\ and\ B = X^T$
         1. $trace(AB) = trace(BA) => trace(X(X^T X)^{-1} X^T) =$
            $trace(XX^T (X^T X)^{-1})\ which\ is\ equal\ to\ trace\left(I_{(d+1)*(d+1)}\right) =$
            $d + 1$
         2. $So\ trace(H) = d + 1$