Rohan Nayar
rn6779

EE 379K Lab 7 Written Questions

6) a) least squares optimization problem with data X and y:

$$\min_{\beta} : \|X\beta - y\|_2^2 = \Sigma(x_i\beta - y_i)^2$$

$$\|X\beta - y\|_2^2 = (X\beta - y)^T(X\beta - y) = X^T\beta^T X\beta - X^T\beta^T y - y^T X\beta + y^T y$$

note that: $(\beta^T X^T y)^T = (y^T X\beta)$ is a 1x1 matrix, so $\beta^T X^T y = y^T X\beta$ and we can simplify to

$$\rightarrow y^T y - 2\beta^T X^T y + \beta^T X^T X\beta$$

differentiate wrt β and set to 0

$$\rightarrow -X^T y + (X^T X)\beta = 0$$

$$(X^T X)\beta = X^T y$$

$$\boxed{\hat{\beta}_{LS} = (X^T X)^{-1} X^T y}$$

b) ridge regression problem:

$$\min_{\beta} \|X\beta - y\|_2^2 + \lambda\|\beta\|_2^2 = \Sigma(x_i\beta - y_i)^2 + \lambda\Sigma\beta_i^2$$

$$\|X\beta - y\|_2^2 + \lambda\|\beta\|_2^2 = (X\beta - y)^T(X\beta - y) + \lambda\beta^T\beta$$

$$= X^T\beta^T X\beta - X^T\beta^T y - y^T X\beta + y^T y + \lambda\beta^T\beta$$

$$= y^T y - 2\beta^T X^T y + \beta^T X^T X\beta - \lambda\beta^T\beta$$

derivative

$$\rightarrow -X^T y + (X^T X)\beta + \lambda\beta = 0$$

$$\rightarrow -X^T y + (X^T X + \lambda I)\beta = 0$$

$$\rightarrow (X^T X + \lambda I)\beta = X^T y$$

$$\boxed{\hat{\beta}_R = (X^T X + \lambda I)^{-1} X^T y}$$