REGRESSION& REGULARIZATION EXERCISE

ELIA YAKIN & LITAL BRIDAVSKY

Developing the Ridge coeficiant. Let $L(\cdot)$ be:

$$L(y,\hat{y}) = \sum_{i=1}^{N}\left(y^{(i)}-\hat{y}^{(i)}\right)^2 + \lambda\|w\|_2^2$$

we need to show that $w_{Ridge} = \left(X^TX+\lambda I\right)^{-1}X^Ty$

$$\begin{aligned}
L(y,\hat{y}) &= \sum_{i=1}^{N}\left(y^{(i)}-\hat{y}^{(i)}\right)^2 + \lambda\|w\|_2^2 \\
&= (y-\hat{y})^T(y-\hat{y}) + \lambda w^Tw \\
&= \left(y-X^Tw\right)^T\left(y-X^Tw\right) + \lambda w^Tw \\
&= y^Ty - 2w^TX^Ty + w^TX^TXw + \lambda w^Tw
\end{aligned}$$

Taking the derivative by w and using FOC:

$$\begin{aligned}
\frac{\partial L}{\partial w} &= 0 - 2X^Ty + 2X^TXw + 2\lambda w \underset{FOC}{=} 0 \\
\iff & \left(X^TX+\lambda I\right)w = X^Ty \\
\implies & w_{Ridge} = \left(X^TX+\lambda I\right)^{-1}X^Ty
\end{aligned}$$

Bonus question: Let $G \sim N\left(1,\sigma^2\right)$ and $X' = X\cdot G$ s.t $EX' = E[X\cdot G] = EX$
First notice the followig property about G

$$Var(G) = \sigma^2 = E(G-EG)^2 = EG^2 - E^2G \underset{EG=1}{=} EG^2 - 1 \iff EG^2 = \sigma^2+1$$

Now looking back on the loss of OLS we can write:

$$\begin{aligned}
L &= \left\|y - w^TX'\right\|^2 \\
&= \left(y - Gw^TX\right)^T\left(y-Gw^TX\right) \\
&= \left[y^Ty - 2Gw^TX^Ty + G^2w^TX^TXw\right] \\
&= \left[y^Ty - G2w^TX^Ty + G^2w^TX^TXw\right]
\end{aligned}$$

$$\begin{aligned}
EL &= E\left[y^Ty - G2w^TX^Ty + G^2w^TX^TXw\right] \\
&= E\left[y^Ty\right] - 2EGEw^TX^Ty + EG^2E\left[w^TX^TXw\right] \\
&= E\left[y^Ty\right] - 2Ew^TX^Ty + \left(1+\sigma^2\right)E\left[w^TX^TXw\right] \\
&= MSE(L) + \sigma^2E\left[w^TX^TXw\right] \\
&= MSE(L) + \sigma^2w^TE\left[X^TX\right]w \\
&= MSE(L) + \sigma^2\cdot\sigma_X^2w^Tw \\
&\underset{*}{=} MSE(L) + \sigma^2\cdot w^Tw
\end{aligned}$$

* if X is noramilized and $\forall l,m\in\{1,2,\dots,p\}\ \ s.t\ l\neq m\ Cov(x_l,x_m)=0$ then $\sigma_X^2 = I_p$ and $X\sim N(0,I_p)$

From the above we can see that we get the Ridge Loss with $\lambda=\sigma^2$