

1. **Please write the formulation for the least square regression, the ridge regression, the kernel regression, and the LASSO regression.**

least square regression:

$$\|Y - X^\top \beta\|^2$$

ridge regression:

$$\|Y - X^\top \beta\|^2 + \lambda \|\beta\|_2^2$$

kernel regression:

$$\|Y - Kc\|^2 + \lambda c^\top Kc$$

LASSO regression:

$$\frac{1}{2} \|Y - X^\top \beta\|^2 + \lambda |\beta|_1$$

2. **Please write analytic solutions to the least square regression, the ridge regression, the kernel regression, and the spline regression. (不用写证明过程)**

least square regression:

$$\beta = (X^\top X)^{-1} X^\top Y$$

ridge regression:

$$\beta = (X^\top X + \lambda I)^{-1} X^\top Y$$

kernel regression:

$$c = (K + \lambda I)^{-1} Y$$

spline regression:

$$\alpha = (Z^\top Z + \lambda D)^{-1} Z^\top Y$$

3. **Please derive the LDA for two classes.**

$$\begin{aligned} \forall X_i \in \Omega^+, p(X_i|y = +1) &\sim N(\mu^+, \mu^-) \\ \forall X_i \in \Omega^-, p(X_i|y = -1) &\sim N(\mu^-, \mu^+) \end{aligned}$$

$$\begin{aligned}
\sigma_{\text{between}}^2 &= [(\mu^+ - \mu^-)^\top \beta]^2 \\
\sigma_{\text{within}}^2 &= n_{\text{pos}} \sigma_{\text{pos}}^2 + n_{\text{neg}} \sigma_{\text{neg}}^2 \\
n_{\text{pos}} &= |\Omega^+| \\
n_{\text{neg}} &= |\Omega^-| \\
\sigma_{\text{pos}}^2 &= \beta^\top \Sigma^+ \beta \\
\sigma_{\text{neg}}^2 &= \beta^\top \Sigma^- \beta
\end{aligned}$$

$$\begin{aligned}
S &= \frac{\sigma_{\text{between}}^2}{\sigma_{\text{within}}^2} \\
&= \frac{[(\mu^+ - \mu^-)^\top \beta]^2}{n_{\text{pos}} \sigma_{\text{pos}}^2 + n_{\text{neg}} \sigma_{\text{neg}}^2} \\
&= \frac{\beta^\top S_B \beta}{\beta^\top S_W \beta}
\end{aligned}$$

其中

$$\begin{aligned}
S_B &= (\mu^+ - \mu^-) (\mu^+ - \mu^-)^\top \\
S_W &= n_{\text{pos}} \Sigma^+ + n_{\text{neg}} \Sigma^-
\end{aligned}$$

因为 β 模长不改变结果，可令

$$\beta^\top S_W \beta = 1$$

$$\max_{\beta} \beta^\top S_B \beta \quad \text{s.t.} \quad \beta^\top S_W \beta = 1$$

$$\begin{aligned}
L &= \beta^\top S_B \beta - \lambda (\beta^\top S_W \beta - 1) \\
\Rightarrow \frac{\partial L}{\partial \beta} &= 2S_B \beta - 2\lambda S_W \beta = 0 \\
\Rightarrow S_B \beta &= \lambda S_W \beta \\
\Rightarrow S_W^{-1} S_B \beta &= \lambda \beta
\end{aligned}$$

β 是 $S_W^{-1} S_B$ 的特征向量。

因为 $S_B \beta$ 与 $\mu^+ - \mu^-$ 方向相同，

$$\beta \propto S_W^{-1} (\mu^+ - \mu^-)$$