Machine Learning – Assignment 6

103062528 林玉山

1. 證明 Semiparametric Representation Theorem 可以套用在 LapRLS 跟 LapSVM 上

首先回顧一下 Semiparametric Representation Theorem 的證明。該證明的方向主要分成兩步，第一步是證明 predict function g 可以分成兩個分量，分別是垂直與平行於 $\mathrm{span}\left(k\left(x^{(1)},\cdot\right), k\left(x^{(2)},\cdot\right) \dots, k\left(x^{(N)},\cdot\right)\right)$ (以下稱為 K 空間) 的向量。接著證明垂直於 K 空間的向量並不會影響 minimization 的結果。因此可以只留平行的向量即可。第二步是證明 RKHS 的 Normalization 同樣在上述情況中不會被影響，因此也可以省略掉。

在這次的問題中，主要是加入了 Laplacian Term。因此只要能夠證明 g 在垂直於 K 空間的向量，並不會對 Laplacian Term 在 Minimization 造成負面影響，就可以套用 Semiparametric Representation Theorem。

(待續)

2. 解出下式中的 $\boldsymbol{\alpha}$

(a) Kernel RLS

以下為套用 Representer Theorem 之後的 Kernel RLS Objective

$$\arg\ \min_{\alpha} \|r - K\alpha\|^2 + \lambda\alpha^\top K\alpha$$

我們拿這個 objective 對 α 微分：

$$\frac{\partial\ (r - K\alpha)^\top (r - K\alpha) + \lambda\alpha^\top K\alpha}{\partial\alpha}$$
$$= -2K^\top (r - K\alpha) + (K + K^\top)\alpha$$

又因為 K 是 Symmetric Matrix，所以 $K = K^\top$

$$= -2K(r - K\alpha) + 2K\alpha$$
$$= 2K(\alpha - r + K\alpha)$$

而此 Objective 的最小值會出現在其導數為 0 之時，因此：

$$2K(\alpha - r + K\alpha) = 0$$
$$\Rightarrow \alpha - r + K\alpha = 0$$

接著再解出 α

$$\Rightarrow \alpha(I + K) = r$$
$$\Rightarrow \alpha(I + K)(I + K)^{-1} = r(I + K)^{-1}$$
$$\Rightarrow \alpha = r(I + K)^{-1}$$

因此 $\alpha = r(I + K)^{-1}$

(b) LapRLS

以下為套用 Semiparametric Representer Theorem 之後的 LapRLS Objective

$$\arg\min_{\alpha}(r-K\alpha)^T A(r-K\alpha)+\mu(K\alpha)^\top LK\alpha+\lambda\alpha^\top K\alpha$$

$$A_{i,i}=1, \text{for i} = 1\sim\text{N}, \text{otherwise}, 0$$

我們拿這個 objective 對 α 微分：

$$\frac{\partial\ (r-K\alpha)^T A(r-K\alpha)+\mu(K\alpha)^\top LK\alpha+\lambda\alpha^\top K\alpha}{\partial\alpha}$$
$$=-K^\top A^\top(r-K\alpha)-K^\top A(r-K\alpha)+\mu(-K^\top K\alpha-K^\top K\alpha)+(K+K^\top)\alpha$$

又因為 A 跟 K 是 Symmetric Matrix，所以 $K=K^\top$ 且 $A=A^\top$

$$=-KA(r-K\alpha)-KA(r-K\alpha)+\mu(-KK\alpha-KK\alpha)+(K+K)\alpha$$
$$=-2KA(r-K\alpha)-2\mu KK\alpha+2K\alpha$$

而此 Objective 的最小值會出現在其導數為 0 之時，因此：

$$-2KA(r-K\alpha)-2\mu KK\alpha+2K\alpha=0$$
$$\Rightarrow -KAr-KAK\alpha-\mu KK\alpha+K\alpha=0$$
$$\Rightarrow -KAr=KAK\alpha+\mu KK\alpha-K\alpha$$
$$\Rightarrow KAK\alpha+\mu KK\alpha-K\alpha=-KAr$$
$$\Rightarrow (KA+\mu K-I)K\alpha=-KAr$$
$$\Rightarrow (KA+\mu K-I)^{-1}(KA+\mu K-I)K\alpha=(KA+\mu K-I)^{-1}-KAr$$
$$\Rightarrow K\alpha=(KA+\mu K-I)^{-1}-KAr$$
$$\Rightarrow K^{-1}K\alpha=K^{-1}(KA+\mu K-I)^{-1}-KAr$$
$$\Rightarrow \alpha=K^{-1}(KA+\mu K-I)^{-1}-KAr$$

因此 $\alpha=K^{-1}(KA+\mu K-I)^{-1}-KAr$

3. 找出 LapSVM 的 Dual Form

LapSVM 的 Primal Objective 為：

$$\arg\min_{f,\xi} \sum_{t=1}^{L} \xi_t + \mu f^\top L f + \lambda \|f\|_{RKHS}{}^2$$

$$\text{subject to } r^{(t)}\left(f^\top\left(\phi\left(x^{(t)}\right)\right)\right) \geq 1 - \xi_t \text{ and } \xi_t \geq 0, \forall t = 1 \sim L$$

套用 Representer Theorem 之後：

$$\arg\min_{b,c,\xi} \sum_{t=1}^{L} \xi_t + \mu(Kc)^\top LKc + \lambda c^\top Kc$$

$$\text{subject to } r^{(t)}\left(\sum_{j=1}^{N} \mathrm{c}_i \mathrm{K}_{i,j} + b\right) \geq 1 - \xi_t \text{ and } \xi_t \geq 0, \forall t = 1 \sim L$$

轉成 Dual Form 之後：

$$L_g = \sum_{t=1}^{L} \xi_t + \mu(Kc)^\top LKc + \lambda c^\top Kc$$

$$\arg\max_{\alpha,\beta} \inf_{b,c,\xi} + \sum_{t=1}^{L} \alpha_t\left(1 - \xi_t - \left(r^{(t)}\left(\sum_{j=1}^{N} \mathrm{c}_i \mathrm{K}_{i,j} + b\right)\right)\right) + \sum_{t=1}^{L} \beta_t(-\xi_t)$$

$$\text{subject to } \alpha \geq 0, \beta \geq 0$$

此時我們再對 b, c, ξ 作微分，並使其為 0，以取得上式的極限值：

$$\frac{\partial L_g}{\partial b} = \sum_{t=1}^{L} \alpha_t r^{(t)} = 0$$

$$\frac{\partial L_g}{\partial c} = \mu(-K^\top Kc - K^\top Kc) + 2\lambda Kc - \sum_{t=1}^{L} \alpha_t r^{(t)} \sum_{j=1}^{N} \mathrm{K}_{i,j} = 0$$

$$\frac{\partial L_g}{\partial \xi_t} = 1 - \alpha_t - \beta_t = 0$$

其中我們可以已從第一個與第三個結果重新改寫 Dual Form

$$L_g = \mu(Kc)^\top LKc + \lambda c^\top Kc + \sum_{t=1}^{L} \alpha_t\left(1 - \xi_t - \left(r^{(t)}\left(\sum_{j=1}^{N} \mathrm{c}_i \mathrm{K}_{i,j} + b\right)\right)\right)$$

$$\Rightarrow \mu c^{\top}K^{\top}LKc + \lambda c^{\top}Kc + \sum_{t=1}^{L} \alpha_t\left(1-\left(r^{(t)}\left(\sum_{j=1}^{N} \mathrm{c}_i \mathrm{K}_{i,j} + b\right)\right)\right)$$

此時我們定義一個矩陣 A=[I, 0], $I \in R^{L,L}, A \in R^{L,N}$，代入上式

$$L_g = \mu c^{\top}K^{\top}LKc + \lambda c^{\top}Kc + \sum_{t=1}^{L} \alpha_t - c^{\top}KA^{\top}R\alpha$$

$$\Rightarrow c^{\top}(\mu K^{\top}LK + \lambda K)c + \sum_{t=1}^{L} \alpha_t - c^{\top}KA^{\top}R\alpha$$

剛剛上面嘗試對 c 偏微分不太好解，這邊再度嘗試一次

$$\frac{\partial L_g}{\partial c} = 2(\mu K^{\top}LK + \lambda K)c - KA^{\top}R\alpha = 0$$

$$\Rightarrow 2(\mu K^{\top}LK + \lambda K)c = KA^{\top}R\alpha$$

$$\Rightarrow (\mu K^{\top}LK + \lambda K)^{-1}2(\mu K^{\top}LK + \lambda K)c = (\mu K^{\top}LK + \lambda K)^{-1}KA^{\top}R\alpha$$

$$\Rightarrow 2c = (\mu K^{\top}LK + \lambda K)^{-1}KA^{\top}R\alpha$$

$$\Rightarrow 2c = (\mu K^{\top}L + \lambda)^{-1}K^{-1}KA^{\top}R\alpha$$

$$\Rightarrow 2c = (\mu K^{\top}L + \lambda)^{-1}A^{\top}R\alpha$$

$$\Rightarrow c = \frac{1}{2}(\mu K^{\top}L + \lambda)^{-1}A^{\top}R\alpha$$

此時再將 c 帶回 L_g

$$L_g = \left(\frac{1}{2}(\mu K^{\top}L + \lambda)^{-1}A^{\top}R\alpha\right)^{\top}(\mu K^{\top}LK + \lambda K)\frac{1}{2}(\mu K^{\top}L + \lambda)^{-1}A^{\top}R\alpha + \sum_{t=1}^{L} \alpha_t - \left(\frac{1}{2}(\mu K^{\top}L + \lambda)^{-1}A^{\top}R\alpha\right)^{\top}KA^{\top}R\alpha$$

$$\Rightarrow \frac{1}{4}((\mu K^{\top}L + \lambda)^{-1}A^{\top}R\alpha)^{\top}KA^{\top}R\alpha + \sum_{t=1}^{L} \alpha_t - \frac{1}{2}((\mu K^{\top}L + \lambda)^{-1}A^{\top}R\alpha)^{\top}KA^{\top}R\alpha$$

$$\Rightarrow \sum_{t=1}^{L} \alpha_t - \frac{1}{4}((\mu K^{\top}L + \lambda)^{-1}A^{\top}R\alpha)^{\top}KA^{\top}R\alpha$$

$$\Rightarrow \sum_{t=1}^{L} \alpha_t - \frac{1}{4}\alpha^{\top}((\mu K^{\top}L + \lambda)^{-1}A^{\top}R)^{\top}KA^{\top}R\alpha$$

令 $Q = ((\mu K^{\top}L + \lambda)^{-1}A^{\top}R)^{\top}KA^{\top}R$

$$\Rightarrow \sum_{t=1}^{L} \alpha_t - \frac{1}{4}\alpha^{\top}\mathrm{Q}\alpha \ subject\ to\ \sum_{t=1}^{L} \alpha_t r^{(t)} = 0, 0 \le \alpha_t \le 1, t = 1\sim L$$

4. 實作 Kernel RLS 與 LapRLS

請參考 README.pdf

5, 6 不能亡

好難喔 OAQ