Carpeta 2, Cuaderno 1

Ejercicio 1 Presente la solución al problema de optimización Kernel Ridge

Sean: $\mathbb{Y} \in \mathbb{R}^{N}$; $\mathbb{X} \in \mathbb{R}^{N \times P}$; $\varphi : \mathbb{R}^{P} \to H$; $H \subseteq \mathbb{R}^{Q}$; $Q \to \infty$

y, $\hat{\mathbb{Y}} = f(\varphi(\mathbb{X}) | \mathbb{W}) = \phi(\mathbb{X})\mathbb{W}$; $f : \mathbb{R}^{Q} \to \mathbb{R}$; $\mathbb{W} \in H \subseteq \mathbb{R}^{Q}$

El problema de optimización tiene la forma:

$$\mathbb{W}^{*} = \underset{\mathbb{W}}{\arg\min}\ L(\mathbb{Y}, f(\varphi(\mathbb{X}) | \mathbb{W})) + R(f | \lambda)$$

$$= \underset{\mathbb{W}}{\arg\min}\ \frac{1}{N} \|\mathbb{Y} - \phi(\mathbb{X})\mathbb{W}\|_2^2 + \lambda \|\mathbb{W}\|_2^2$$

Se tiene,

$\lambda \in \mathbb{R}^{+}$: Parametro de regularizacion, Por ridge, $\lambda \to$ fijo

Se expande el termino cuadratico,

$$\|\mathbb{Y} - \phi(\mathbb{X})\mathbb{W}\|_2^2 = \langle \mathbb{Y} - \phi(\mathbb{X})\mathbb{W}, \mathbb{Y} - \phi(\mathbb{X})\mathbb{W} \rangle$$

$$= \mathbb{Y}^T \mathbb{Y} - 2\mathbb{Y}^T \phi(\mathbb{X})\mathbb{W} + (\phi(\mathbb{X})\mathbb{W})^T (\phi(\mathbb{X})\mathbb{W} >$$

$$= \mathbb{Y}^T \mathbb{Y} - 2\mathbb{Y}^T \phi(\mathbb{X})\mathbb{W} + \mathbb{W}^T \phi^T(\mathbb{X})\phi(\mathbb{X})\mathbb{W}$$

Nota: $\phi(\mathbb{X})$ es el mapeo de los datos $\mathbb{X}$, $\phi(\mathbb{X}) \in \mathbb{R}^{N \times Q}$

La correlación en RKHS esta dada como,

$$\frac{1}{N}\phi^T(\mathbb{X})\phi(\mathbb{X}) \in \mathbb{R}^{Q \times Q} \simeq E\{\phi(\mathbb{X})\phi(\mathbb{X})\} = R_{\varphi(x)\varphi(x)}$$

Ahora, Se deriva el problema respecto a $\mathbb{W}$

$$\frac{\partial}{\partial \mathbb{W}} \{L(\mathbb{Y}, f) + R(f | \lambda)\}_{Q \times 1} = -\frac{2}{N}\phi^T(\mathbb{X})\mathbb{Y}_{Q \times N \times N \times 1} + \frac{2}{N}\phi^T(\mathbb{X})\phi(\mathbb{X})\mathbb{W}_{Q \times Q \times Q \times 1} + 2\lambda \mathbb{W}_{Q \times 1}$$

$$\nabla \mathbb{W} = \frac{2}{N}[(\phi^T(\mathbb{X})\phi(\mathbb{X}) + N\lambda)\mathbb{W} - \phi^T(\mathbb{X})\mathbb{Y}]$$

Solución analitica

$$\frac{2}{N}[(\phi^T(\mathbb{X})\phi(\mathbb{X}) + N\lambda I)\mathbb{W}] = \frac{2}{N}\phi^T(\mathbb{X})\mathbb{Y}$$

$$\mathbb{W}^{*} = [\phi^T(\mathbb{X})\phi(\mathbb{X}) + N\lambda I]^{-1}\phi^T(\mathbb{X})\mathbb{Y}$$

Teniendo en cuenta,

$$\Phi^T(x)\,\Phi(x) \in \mathbb{R}^{Q \times Q} \; ; \quad Q \to \infty$$

$$(I+AB)^{-1}A = A(I+BA)^{-1}$$ → Propiedades de la matriz inversa

$$(\Phi^T\Phi + N\lambda I)^{-1}\Phi^T = \left[N\lambda\left(\frac{1}{N\lambda}\Phi^T\Phi + I\right)\right]^{-1}\Phi^T$$

$$= \frac{1}{N\lambda}\left(I + \Phi^T\frac{\Phi}{N\lambda}\right)^{-1}\Phi^T = \Phi^T\frac{1}{N\lambda}\left(I + \frac{\Phi}{N\lambda}\Phi^T\right)^{-1}$$

$$= \Phi^T\left[N\lambda\left(I + \frac{\Phi}{N\lambda}\Phi^T\right)\right]^{-1} = \Phi^T\left[N\lambda I + \Phi\Phi^T\right]^{-1}$$

Finalmente, se tiene la solucion como,

$$\boxed{w^* = \Phi^T(x)\left[N\lambda I + \Phi(x)\Phi^T(x)\right]^{-1} y}$$