4. a) $\bar{f}_p = [a(c_1+\bar{x}_p^T\bar{v}_1) \quad a(c_2+\bar{x}_p^T\bar{v}_2) \quad \cdots \quad a(c_M+\bar{x}_p^T\bar{v}_M)]^T$

$$\underset{b,\bar{w},\Theta}{\text{minimize}} \sum_{p=1}^{P} \log(1+e^{-y_p(b+\bar{f}_p\bar{w})}) = g$$

$$\frac{\partial}{\partial b} g = +\sum_{p=1}^{P} \sigma(-y_p(b+\bar{f}_p\bar{w}))(-y_p(b+\bar{f}_p\bar{w}))' = -\sum_{p=1}^{P} \sigma(-y_p(b+\sum_{m=1}^{M} w_m a(c_m+\bar{x}_p^T\bar{v}_m)))\,y_p$$

$$\frac{\partial}{\partial w_n} g = \sum_{p=1}^{P} \sigma(-y_p(b+\sum_{m=1}^{M} w_m a(c_m+\bar{x}_p^T\bar{v}_m))).(-y_p(b+\sum_{m=1}^{M} w_m a(c_m+\bar{x}_p^T\bar{v}_m)))'$$

$$= -\sum_{p=1}^{P} \sigma(-y_p(b+\sum_{m=1}^{M} w_m a(c_m+\bar{x}_p^T\bar{v}_m)))\,a(c_n+\bar{x}_p^T v_n)\,y_p$$

$$\frac{\partial}{\partial c_n} g = \sum_{p=1}^{P} \sigma(-y_p(b+\sum_{m=1}^{M} w_m a(c_m+\bar{x}_p^T\bar{v}_m)))\,(-y_p w_n a(c_n+\bar{x}_p^T\bar{v}_n))'$$

$$= -\sum_{p=1}^{P} \sigma(-y_p(b+\sum_{m=1}^{M} w_m a(c_m+\bar{x}_p^T\bar{v}_m)))\,a'(c_n+\bar{x}_p^T\bar{v}_n)w_n y_p$$

$$\nabla_{\bar{v}_n} g = \sum_{p=1}^{P} \sigma(-y_p(b+\sum_{m=1}^{M} w_m a(c_m+\bar{x}_p^T\bar{v}_m)))(-y_p w_n a(c_n+\bar{x}_p^T\bar{v}_n))'$$

$$= -\sum_{p=1}^{P} (-y_p(b+\sum_{m=1}^{M} w_m a(c_m+\bar{x}_p^T\bar{v}_n)))\,a'(c_n+\bar{x}_p^T\bar{v}_n)\,\bar{x}_p w_n y_p$$

(b) $\tilde{q} = \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_P \end{bmatrix}$ $\quad \tilde{t}_n = \begin{bmatrix} \tanh(c_n+\bar{x}_p\bar{v}_n) \\ \tanh(c_n+\bar{x}_2\bar{v}_n) \\ \vdots \\ \tanh(c_n+\bar{x}_p\bar{v}_n) \end{bmatrix}$ $\quad \tilde{S}_n = \begin{bmatrix} \operatorname{sech}^2(c_n+\bar{x}_1^T\bar{v}_n) \\ \operatorname{sech}^2(c_n+\bar{x}_2^T\bar{v}_n) \\ \vdots \\ \operatorname{sech}^2(c_n+\bar{x}_n^T\bar{v}_n) \end{bmatrix}$

$\tilde{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{bmatrix}$ $\quad \tilde{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}$

$$\frac{\partial}{\partial b} g = -\tilde{1}_{p\times 1}^T\,\tilde{q}\odot\tilde{y}.$$

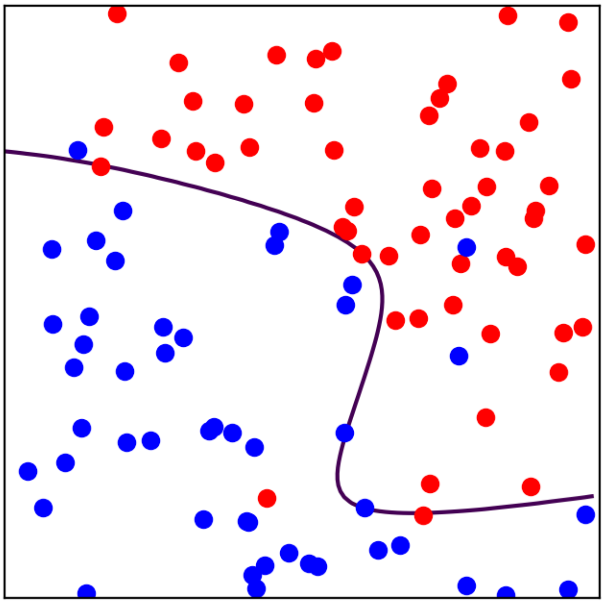
$$\frac{\partial}{\partial w_n} g = -\tilde{1}_{p\times 1}^T(\tilde{q}\odot\tilde{t}_n\odot\tilde{y}).$$

$$\frac{\partial}{\partial c_n} g = -\tilde{1}_{p\times 1}^T(\tilde{q}\odot\tilde{S}_n\odot\tilde{y})\,w_n.$$

$$\nabla_{\bar{v}_n} g = -\tilde{X}\cdot\tilde{q}\odot\tilde{S}_n\odot\tilde{y}\,w_n$$

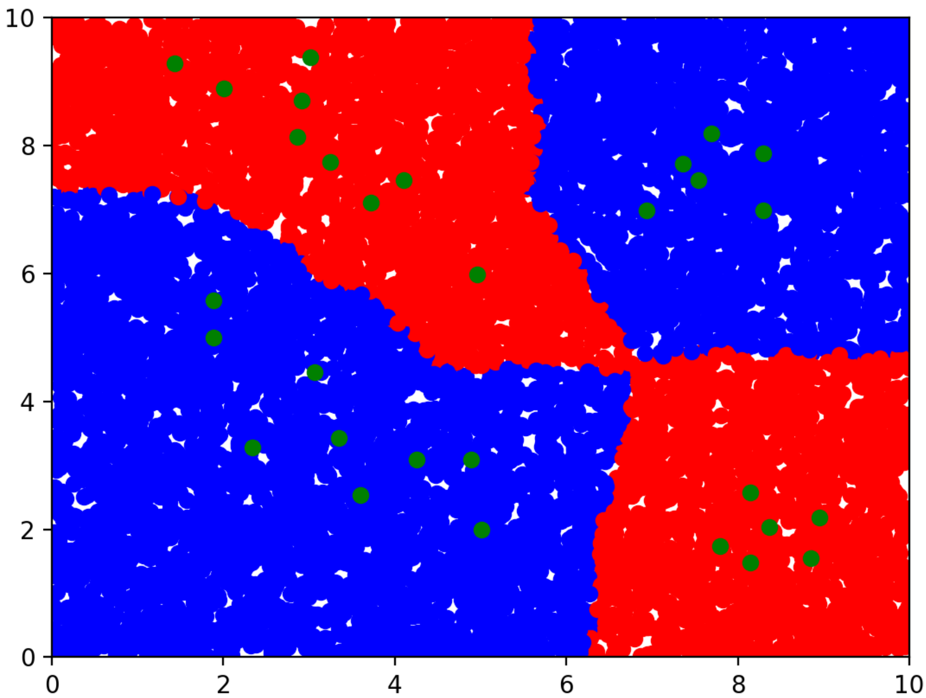
由 扫描全能王 扫描创建

6.5

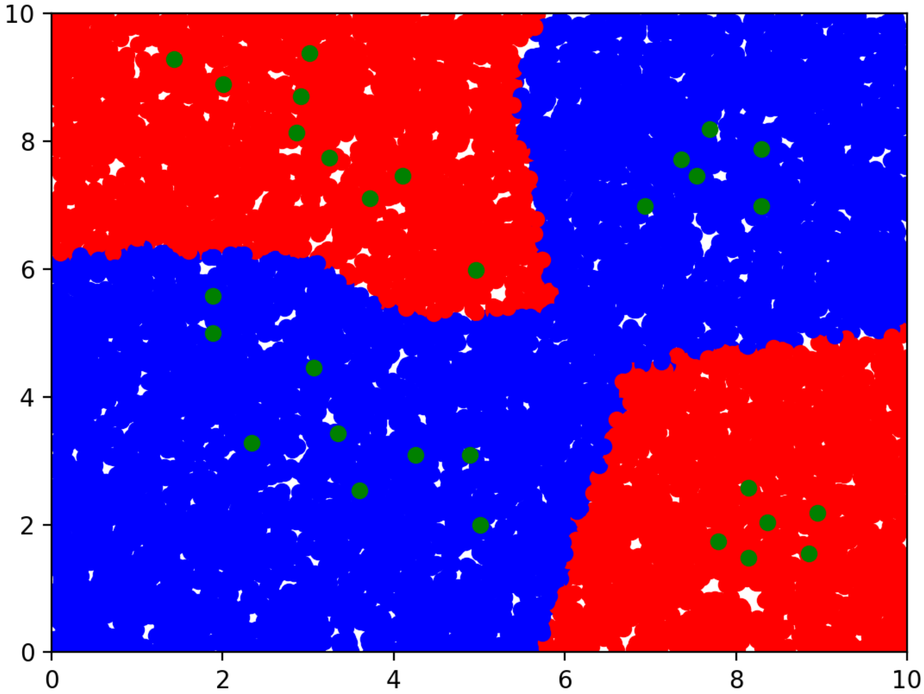


6.6

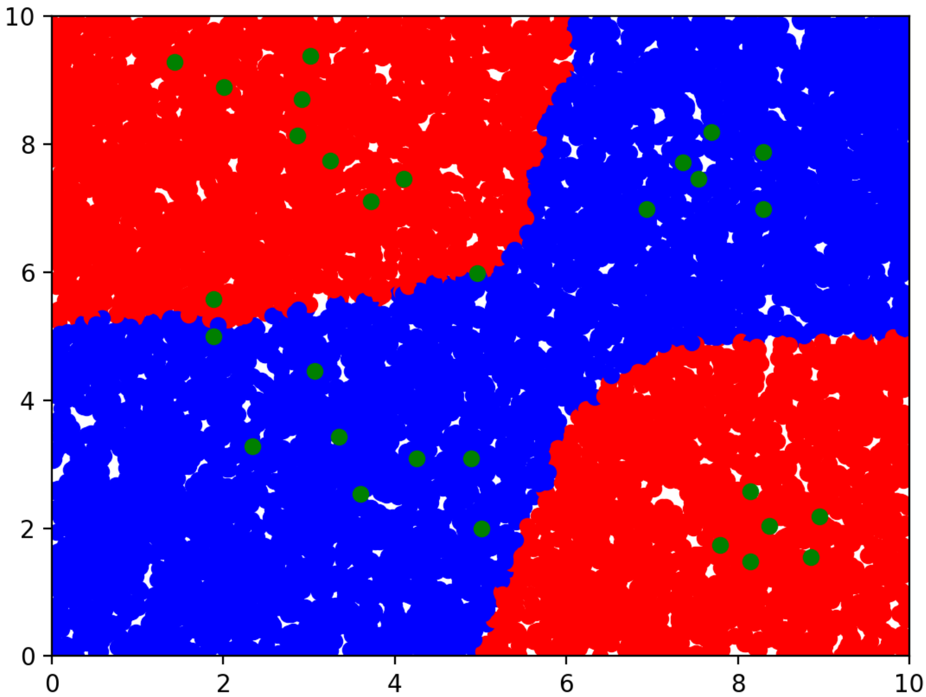
k = 1



k = 5



k = 10



6.9 According to the figure, the best degree of polynomial basis is 5.