

付録

1 KKT 最適条件を用いたラグランジュ関数の変形

$$L(\mathbf{w}, \gamma, \boldsymbol{\xi}, \boldsymbol{\alpha}, \boldsymbol{\beta}) = \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n \xi_i - \sum_{i=1}^n \alpha_i \{(\mathbf{w}^T \mathbf{x}_i + \gamma) y_i - 1 + \xi_i\} - \sum_{i=1}^n \beta_i \xi_i \quad (1)$$

について、

$$\frac{\partial L}{\partial \mathbf{w}} = \begin{pmatrix} \frac{\partial L}{\partial w_1} \\ \vdots \\ \frac{\partial L}{\partial w_n} \end{pmatrix} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} - \begin{pmatrix} \sum_{i=1}^n \alpha_i x_i^{(1)} y_i \\ \vdots \\ \sum_{i=1}^n \alpha_i x_i^{(n)} y_i \end{pmatrix} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} - \begin{pmatrix} \sum_{i=1}^n \alpha_i y_i x_i^{(1)} \\ \vdots \\ \sum_{i=1}^n \alpha_i y_i x_i^{(n)} \end{pmatrix} = \mathbf{w} - \sum_{i=1}^n \alpha_i y_i \mathbf{x}_i = \mathbf{0} \quad (2)$$

$$\frac{\partial L}{\partial \gamma} = \sum_{i=1}^n \alpha_i y_i = 0 \quad (3)$$

$$\frac{\partial L}{\partial \boldsymbol{\xi}} = \begin{pmatrix} C \\ \vdots \\ C \end{pmatrix} - \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} - \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = \mathbf{0} \quad (4)$$

2 ラグランジュ関数における $\mathbf{w}, \boldsymbol{\xi}, \gamma$ の消去

$$\mathbf{w} = \sum_{i=1}^n \alpha_i y_i \mathbf{x}_i, \quad \boldsymbol{\alpha} + \boldsymbol{\beta} = \mathbf{C}$$

により、

$$\begin{aligned} L(\mathbf{w}, \gamma, \boldsymbol{\xi}, \boldsymbol{\alpha}, \boldsymbol{\beta}) &= \frac{1}{2} (\sum_{i=1}^n \alpha_i y_i \mathbf{x}_i)^T (\sum_{j=1}^n \alpha_j y_j \mathbf{x}_j) + \sum_{i=1}^n (\alpha_i + \beta_i) \xi_i - \sum_{i=1}^n \alpha_i ((\sum_{j=1}^n \alpha_j y_j \mathbf{x}_j)^T \mathbf{x}_i + \gamma) y_i - 1 + \xi_i - \sum_{i=1}^n \beta_i \xi_i \\ &= \frac{1}{2} (\sum_{i,j=1}^n \alpha_i y_i \alpha_j y_j \mathbf{x}_i^T \mathbf{x}_j) + \sum_{i=1}^n (\alpha_i + \beta_i) \xi_i - \sum_{i=1}^n \alpha_i \sum_{j=1}^n \alpha_j y_j \mathbf{x}_j^T \mathbf{x}_i y_i + \gamma \sum_{i=1}^n \alpha_i y_i - \sum_{i=1}^n \alpha_i + \sum_{i=1}^n \alpha_i \xi_i - \sum_{i=1}^n \beta_i \xi_i \\ &= \frac{1}{2} (\sum_{i,j=1}^n \alpha_i y_i \alpha_j y_j \mathbf{x}_i^T \mathbf{x}_j) + \sum_{i=1}^n (\alpha_i + \beta_i) \xi_i - \sum_{i,j=1}^n \alpha_i \alpha_j y_j \mathbf{x}_i^T \mathbf{x}_j y_i + \gamma \sum_{i=1}^n \alpha_i y_i - \sum_{i=1}^n \alpha_i + \sum_{i=1}^n \alpha_i \xi_i - \sum_{i=1}^n \beta_i \xi_i \\ &= \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j \mathbf{x}_i^T \mathbf{x}_j \quad (5) \end{aligned}$$