Lagrangian equation:

$$L(b, w, \lambda) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{N} \lambda_i [y_i (w^T x_i + b) - 1] \quad \text{—} (1)$$

Let me define X, M matrix as follows:

$$X = \begin{bmatrix} \text{——} x_1 \text{——} \\ \text{——} x_2 \text{——} \\ \vdots \\ \text{——} x_N \text{——} \end{bmatrix}_{N \times p}$$

where $p$ = no. of features in data point. & $x_1, x_2 \cdots x_N$ are my data points.

$$M = \begin{bmatrix} \text{——} x_1 y_1 \text{——} \\ \text{——} x_2 y_1 \text{——} \\ \vdots \\ \text{——} x_N y_N \text{——} \end{bmatrix}_{N \times p}$$

multiplying each row of x with corresponding elements of Y.

where $Y = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}_{N \times 1}$, Assume $\mathring{c} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}_{N \times 1}$

Let $\lambda = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_N \end{bmatrix}_{N \times 1} \Rightarrow \lambda^T = [\lambda_1, \cdots, \lambda_n]_{1 \times N}$

So I can write

$\sum_{i=1}^{N} \lambda_i [y_i (w^T x_i + b) - 1]$ in matrix form as:

$\lambda^T [Mw + bY - \mathring{c}]$ where corresponding matrices are defined as above.

so Lagrangian can be rewritten as:-

$$L(b,w,\lambda) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{N} \lambda_i (y_i(w^T x_i + b) - 1)$$

$$= \frac{1}{2} w^T w - \lambda^T [Mw + by - i] \quad — (2)$$

Deriving KKT condition:

$$\nabla_w L = w - M^T \lambda = 0 \Rightarrow \boxed{w = M^T \lambda}$$

$$\Rightarrow \boxed{w = \sum_{i=1}^{N} \lambda_i y_i x_i} \quad — (3)$$

$$\nabla_b L = -\lambda^T Y = 0 \Rightarrow \boxed{\lambda^T Y = 0}$$

$$\Rightarrow \boxed{\sum_{i=1}^{N} \lambda_i y_i = 0} \quad — (4)$$

using (3) & (4) in (2) we get

$$L(b,w,\lambda) = \frac{1}{2} w^T w - \lambda^T (Mw + b\cancel{y} - i]$$

Replacing $w = M^T \lambda$ & $\lambda^T Y = 0$

$$= \frac{1}{2} \lambda^T M M^T \lambda - \lambda^T [M M^T \lambda + by - i]$$

$$= \frac{1}{2} \lambda^T M M^T \lambda - \lambda^T M M^T \lambda + \lambda^T i$$

Thus we can write dual function $g(\lambda)$ as:

$$g(\lambda) = \begin{cases} -\frac{1}{2} \lambda^T M M^T \lambda + \lambda^T I & \text{if } \lambda^T Y = 0 \\ -\infty & \text{Otherwise} \end{cases}$$

we can open $mm^T$ as follow.

$$m = \begin{bmatrix} - x_1 y_1 - \\ \vdots \\ - x_N y_N - \end{bmatrix} \Rightarrow mm^T = \begin{bmatrix} y_1 y_1 x_1 x_1^T & \cdots & y_1 y_N x_1 x_N^T \\ \vdots & & \\ y_N y_1 x_N x_1^T & \cdots & y_N y_N x_N x_N^T \end{bmatrix}$$

where $x_i$ is a row in $X$ vector (data point)

Thus we can write

$$g(\lambda) = \begin{cases} -\frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} y_i y_j \lambda_i \lambda_j x_i x_j^T + \sum_{i=1}^{N} \lambda_i & \text{if } \sum_{i=1}^{N} \lambda_i y_i = 0 \\ -\infty & \text{otherwise} \end{cases}$$

Thus $\boxed{g(\lambda) = \Lambda^T B + \frac{1}{2} \Lambda^T A \Lambda}$

where $\Lambda = \lambda = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix}$ $\quad B = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}_{N \times 1} = \mathbb{1}$

$$\boxed{A = -mm^T}$$