4.1.7 Maximize: $f(x,y) = 16 - (x^2+y^2)$.

subject to : $2x - y + 4 = 0$.

= $L(x,y,\lambda) = 16 - (x^2+y^2) + \lambda(2x - y + 4)$

= Maximize over x & y:

① $\frac{\partial L}{\partial x} = -2x + 2\lambda = 0 \Rightarrow \boxed{\lambda = x}$

② $\frac{\partial L}{\partial y} = -2y - \lambda = 0 \Rightarrow \boxed{y = -\lambda/2}$

= substituting back in L ⇒ dual function.

$$L(\lambda) = 16 - \left(\lambda^2 + \frac{\lambda^2}{4}\right) + \lambda\left(2\lambda + \frac{\lambda}{2} + 4\right)$$

$$= 16 - \frac{5\lambda^2}{4} + \frac{5\lambda^2}{2} + 4\lambda = 16 + \frac{5\lambda^2}{4} + 4\lambda$$

= minimize the dual function $L(\lambda)$

= $\frac{dL}{d\lambda} = \frac{5\lambda}{2} + 4 \rightarrow \boxed{\lambda = -8/5}$, $\frac{d^2L}{d\lambda^2} = \frac{5}{2} > 0$ (minima) ✓

= solution:

$$x = \lambda = -8/5, \quad y = -\lambda/2 = 4/5$$

= verification:

① $2x - y + 4 = 0 \Rightarrow 2(-8/5) - 4/5 + 4 = 0$ ✓

② $16 - (x^2+y^2) \Rightarrow 16 - (64/25 + 16/25) = 16 - 80/25 = 64/5 = 12.8$

4.4.7 Starting from equation in (4.17)

$$= L(w_1, w_2, b, \lambda) = \frac{w_1^2 + w_2^2}{2} + \sum_{i=1}^{n} \lambda_i (1 - z_i (w_1 x_i + w_2 y_i + b))$$

= From the partial derivative equations (4.18), we have:

- $w_1 = \sum_i^n \lambda_i . z_i . x_i$
- $w_2 = \sum_i^n . \lambda_i . z_i . y_i$
- $\sum_i^n \lambda_i . z_i = 0$

= This gives us $w = \sum_i^n \lambda_i . z_i . X_i$ where $X_i = (x_i, y_i)$

= Substituting this in the Lagrangian:

$$L(\lambda) = \frac{1}{2}\left(\sum_1^n \lambda_i . z_i . x_i\right)^2 + \frac{1}{2}\left(\sum_1^n \lambda_i . z_i . y_i\right)^2 + \sum_{i=1}^n \lambda_i - \sum_{i=1}^n \lambda_i . z_i (w_1 . x_i + w_2 y_i + b)$$

= since $\sum^n \lambda_i . z_i = 0$, the term with b vanishes. Quadratic terms combine as:

$$\frac{1}{2}\left(\sum_1^n \lambda_i . z_i . x_i\right)^2 + \frac{1}{2}\left(\sum_1^n \lambda_i . z_i . y_i\right)^2 = \frac{1}{2} \sum_i^n \sum_j^n \lambda_i . \lambda_j . z_i . z_j (x_i x_j + y_i y_j)$$

$$\boxed{L(\lambda) = \sum_{i=1}^n \lambda_i - \frac{1}{2} \sum_i^n \sum_j^n \lambda_i . \lambda_j . z_i . z_j (X_i . X_j)}$$

~~4.12.~~ 4.8.) (a) For a linear SVM, the scoring function from equation 4.22

$= \quad f(x) = \sum_{i}^{n} \lambda_i . z_i \, (x_i . x) + b.$

$=$ since $x_i . x = x_i x + y_i y$ for $x = (x, y)$ and $x_i = (x_i, y_i)$,

$= \quad f(x) = \left(\sum_{i}^{n} \lambda_i . z_i . x_i \right) x + \left(\sum_{i}^{n} \lambda_i . z_i . y_i \right) y + b.$

$=$ ∴ the weights are

$x : w_1 = \sum_{i}^{n} \lambda_i . z_i . x_i.$

$y : w_2 = \sum_{i}^{n} \lambda_i . z_i . y_i.$

(b) To reduce dimensionality using SVM weights:

i.) Features with larger $|w_i|$ values contribute more.

ii.) Eliminate features with weight below a threshold

iii.) The SVM can be retrained on feature set.