

$$\text{SVM: } \max_{\alpha_i \in [0,1], i=1..n} \sum \alpha_i - \frac{1}{2\lambda} \alpha^T Y A A^T Y \alpha$$

$$w = \frac{1}{\lambda} A^T Y \alpha.$$

Shapes: $\alpha: () \ n \times 1$

$Y: (\backslash) \ n \times n$ diagonal

$A: () \ n \times d$

$A A^T: () \ n \times n$ (similarity in features)

$Y A A^T Y$: similarities $\times 1$ if same label, $\times -1$ if not.
symmetric.

Optimize 1 coordinate. Projected GD

$$0 = \frac{\partial}{\partial \alpha_i} \text{SVM} = 1 - \frac{1}{\lambda} [Y A A^T Y]_{i,i} \alpha_i$$

$$= 1 - \frac{1}{\lambda} [Y A A^T Y]_{i,i} \alpha_i - \frac{1}{\lambda} [Y A^T A Y]_{i,-i} \alpha_{-i}$$

$$\frac{1}{\lambda} [A A^T]_{i,i} \alpha_i$$

$$\frac{1}{\lambda} A_i^T A_i \alpha_i = \frac{1}{\lambda} \|A_i\|^2 \alpha_i$$

$$\alpha_i = \frac{\lambda - [Y A A^T Y]_{i,i} \alpha_i}{\|A_i\|^2} = \frac{\lambda - [Y A A^T Y]_{i,i} \alpha_i + \|A_i\|^2 \alpha_i}{\|A_i\|^2}$$

$$= \alpha_i + \frac{\lambda}{\|A_i\|^2} - \frac{Y_i A_i^T A^T Y \alpha}{\|A_i\|^2}$$

$$= \alpha_i + \lambda \cdot \frac{1 - Y_i A_i^T w}{\|A_i\|^2}$$

Link to notes on SVM.
lecture 7 from ML course

$$\begin{pmatrix} \otimes \end{pmatrix} \begin{pmatrix} \otimes \end{pmatrix}$$

$$\begin{pmatrix} * \end{pmatrix} \begin{pmatrix} - \end{pmatrix} \begin{pmatrix} | \end{pmatrix}$$

ith row, as a row