Common application: in important cases, one may be able to write the dual directly.

SVM

$$\min \frac{w'w}{2} \quad \text{st } y_i (w'x_i + b) \geq 1$$

Primal form, Separable

$$\mathcal{L}_P(w, \lambda) = \frac{w'w}{2} - \sum_i \lambda_i \left\{ \left[y_i (w'x_i + b) \right] - 1 \right\}$$

$$\nabla_w \mathcal{L} = 0 = w - \sum_i \lambda_i \left\{ \left[y_i x_i \right] \right\}$$

$$\nabla_b \mathcal{L} = 0 = - \sum_i \lambda_i y_i$$

Subst

$$\mathcal{L}_D = \sum_i \lambda_i - \frac{1}{2} \sum_{ij} \lambda_i \lambda_j \left[y_i y_j (x_i^T x_j) \right]$$

Notice constraints

$$\sum_i \lambda_i y_i = 0$$

$$\lambda_i \geq 0$$

and we must max this in λ

If there is an fp for primal, the max is soln to primal

i.e. Value (Dual) = Value (Primal)

What if data is not separable?

$$\min \frac{w'w}{2} + C\sum_i \xi_i$$

$$\text{st} \quad y_i(w'x_i + b) \geq 1 - \xi_i$$

$$\xi_i \geq 0$$

Primal prob

ξ_i are <u>slack variables</u>

$$\mathcal{L}_p = \frac{w'w}{2} + C\sum_i \xi_i - \sum_i \lambda_i [y_i(w'x_i + b) - 1 + \xi_i] - \sum_i \mu_i \xi_i$$

$$\nabla_w \mathcal{L}_p = w - \sum_i \lambda_i y_i x_i = 0$$

$$\nabla_b \mathcal{L}_p = 0 = -\sum_i \lambda_i y_i$$

$$\nabla_{\xi_i} \mathcal{L}_p = C - \lambda_i - \mu_i = 0$$

$\rightarrow$ this gets rid of ξ_i

So we have

$$L_D = \sum_i \lambda_i - \frac{1}{2} \sum_{ij} y_i y_j \lambda_i \lambda_j x_i' x_j$$

subject to

$$\sum_i \lambda_i y_i = 0$$

$$0 \le \lambda_i \le C$$

Notice that ξ_i can be interpreted as a loss

$$\text{hinge loss}(y, y_p) = \max(0, 1 - y_i y_p)$$