Machine Intelligence
Exercise 11
2 Seriously Cool Guys
Due Jan 24, 2018

Primal variables ⇒ $w, b, \varphi_\alpha, \varphi_\alpha^*, \varepsilon$

Dual Variables ⇒ $\lambda_\alpha, \lambda_\alpha^*, \eta_\alpha, \eta_\alpha^*, \delta$

1a.

(1) $\frac{\partial L}{\partial w} = |\underline{w}| + \sum_{\alpha=1}^{p} \lambda_\alpha x_\alpha - \sum_{\alpha=1}^{p} \lambda_\alpha^* x_\alpha$ (2) $\frac{\partial L}{\partial b} = \sum_{\alpha=1}^{p} \lambda_\alpha - \sum_{\alpha=1}^{p} \lambda_\alpha^*$

(3) $\frac{\partial L}{\partial \varphi_\alpha} = \frac{C}{p} - \sum_{\alpha=1}^{p} \eta_\alpha - \sum_{\alpha=1}^{p} \lambda_\alpha$ (4) $\frac{\partial L}{\partial \varphi_\alpha^*} = \frac{C}{p} - \sum_{\alpha=1}^{p} \eta_\alpha^* - \sum_{\alpha=1}^{p} \lambda_\alpha^*$

(5) $\frac{\partial L}{\partial \varepsilon} = C\nu - \sum_{\alpha=1}^{p} \lambda_\alpha - \sum_{\alpha=1}^{p} \lambda_\alpha^* - \delta$

1b.

Setting all derivatives := 0, ..

From (1) we know: $|\underline{w}| = \sum_{\alpha=1}^{p} \lambda_\alpha^* x_\alpha - \lambda_\alpha x_\alpha$

$= \sum_{\alpha=1}^{p} (\lambda_\alpha^* - \lambda_\alpha) x_\alpha$

Given $\eta_\alpha, \eta_\alpha^* \geq 0$ and from (3) & (4), we know: $\frac{C}{p} \geq \sum_{\alpha=1}^{p} \lambda_\alpha$ and $\frac{C}{p} \geq \sum_{\alpha=1}^{p} \lambda_\alpha^*$
our constraints #1 and #2

From (2), we know our constraint #3: $0 = \sum_{\alpha=1}^{p} \lambda_\alpha - \lambda_\alpha^*$

From (5) and given $\delta \geq 0$, we know our constraint #4: $C\nu - \delta = \sum_{\alpha=1}^{p} \lambda_\alpha + \lambda_\alpha^*$

$\Downarrow$

$C\nu \geq \sum_{\alpha=1}^{p} \lambda_\alpha + \lambda_\alpha^*$

Expanding our Lagrangian for workability...

$L = \frac{1}{2}|\underline{w}|^2 + C\nu\varepsilon + \frac{C}{p}\sum_{\alpha=1}^{p} \varphi_\alpha + \varphi_\alpha^* - \sum_{\alpha=1}^{p} \lambda_\alpha \varphi_\alpha - \sum_{\alpha=1}^{p} \lambda_\alpha \varepsilon - \sum_{\alpha=1}^{p} y_T^\alpha \lambda_\alpha$

$+ \sum_{\alpha=1}^{p} \lambda_\alpha w^T x_\alpha + \sum_{\alpha=1}^{p} \lambda_\alpha b - \sum_{\alpha=1}^{p} \lambda_\alpha^* \varphi_\alpha^* - \sum_{\alpha=1}^{p} \lambda^* \varepsilon - \sum_{\alpha=1}^{p} \lambda_\alpha^* y_T^\alpha$

$- \sum_{\alpha=1}^{p} \lambda_\alpha^* w^T x_\alpha - \sum_{\alpha=1}^{p} \lambda_\alpha^* b - \sum_{\alpha=1}^{p} \eta_\alpha \varphi_\alpha - \sum_{\alpha=1}^{p} \eta_\alpha^* \varphi_\alpha^* - \delta\varepsilon$

16. cont'd

$$L = \frac{1}{2}\Big(\overbrace{\sum_{\alpha=1}^{P}(\lambda^*_\alpha - \lambda_\alpha)x_\alpha}^{|\vec{w}|}\Big)^2 + \varepsilon\overbrace{(C\nu - \delta)}^{\sum_{\alpha=1}^{P}\lambda_\alpha + \lambda^*_\alpha} - \sum_{\alpha=1}^{P}\varphi_\alpha\overbrace{(\lambda_\alpha + \eta_\alpha)}^{\frac{C}{P}} - \sum_{\alpha=1}^{P}\varphi^*_\alpha\overbrace{(\lambda^*_\alpha + \eta^*_\alpha)}^{C/P}$$

$$+ \frac{C}{P}\sum_{\alpha=1}^{P}\varphi_\alpha + \varphi^*_\alpha - \varepsilon\sum_{\alpha=1}^{P}(\lambda_\alpha + \lambda^*_\alpha) + \underbrace{\sum_{\alpha=1}^{P}\lambda_\alpha b - \sum_{\alpha=1}^{P}\lambda^*_\alpha b}_{0\text{, given (2)}}$$

$$+ \sum_{\beta,\alpha=1}^{P}\lambda_\alpha\overbrace{(\lambda^*_\beta - \lambda_\beta)x^T_\beta}^{w^T} x_\alpha - \sum_{\beta,\alpha=1}^{P}\lambda^*_\alpha(\lambda^*_\beta - \lambda_\beta)x^T_\beta x_\alpha - \sum_{\alpha=1}^{P}(\lambda_\alpha + \lambda^*_\alpha)y^\alpha_T$$

After applying the square and combining expressions

$$L = -\frac{1}{2}\sum_{\alpha,\beta=1}^{P}(\lambda^*_\alpha - \lambda_\alpha)(\lambda^*_\beta - \lambda_\beta)x^{\alpha T}x_\beta + \sum_{\alpha=1}^{P}(\lambda_\alpha + \lambda^*_\alpha)y^\alpha_T$$

With our Lagrangian, we can now find our solution to the dual problem by finding the maximum & optimal λ_α and λ^*_α.