Machine Learning cheat sheet

General problem

Machine learning problems are almost everytime taking the form:

$$\min_w \frac{1}{n}\sum_{i=1}^{n} l(f_w(x_i), y_i) + \lambda C(w)$$

where:

- l is a loss function
- f_w is the prediction function (w is the parameter to optimize)
- C is a regularization function and λ a regularization factor

Linear case

Here, $f_w(x) = w^T x$, and for a given dataset in $\mathcal{X} \times \mathcal{Y}$, $\phi_i(\cdot) = l(\cdot, y_i)$: We can apply Fenchel duality and we have:

Primal	Dual
$\min_w \frac{1}{n}\sum_{i=1}^{n} \phi_i(w^T x_i) + C(w)$	$\max_\alpha \frac{1}{n}\sum_{i=1}^{n} -\phi_i^*(-\alpha_i) - C^*(\sum_{i=1}^{n} \alpha_i x_i)$

Where ϕ_i^* and C^* are the convex conjugate of ϕ_i and C (defined by $f^*(a) = \max_z(za - f(z))$). And $x_i^*(\cdot)$ is the adjoint operator of $(\cdot^T x_i)$ ($w^{T*} = \overline{w}$)

It can be shown that if $\hat{w}$ and $\hat{\alpha}$ are the optimal solutions of those problems, then we have:

$$P(\hat{w}) = P(w(\hat{\alpha})) = D(\hat{\alpha})$$

Loss functions:

Name	$l(x, y)$	$l^*(-a, y)$
L^1	$\lvert x - y\rvert$	$-ay$, $a \in [-1, 1]$
L^2	$(x - y)^2$	$-ay + \frac{a^2}{4}$
Logistic	$\ln(1 + e^{-xy})$	$ay_i \ln(ay) + (1 - ay)\ln(1 - ay)$, $ay \in [0, 1]$
Hinge loss	$\max(0, 1 - xy)$	$-ay$, $ay \in [0, 1]$

Regularizations:

- L^2: $C(w) = \|\|w\|\|_2^2$
- L^1: $C(w) = \|\|w\|\|_1$

Problems

test

Regression

test

Classification

test