Intro to Statistical Learning Theory
Exercise 5

1) Prove theorem 2.2 in lecture 10.

2) The doubling trick - Our bounds depend on knowing the horizon T (for parameter setting). Let $A(T)$ be an algorithm that runs for T steps and has regret bound $\alpha\sqrt{T}$ for some constant α.

   Consider the algorithm that for $m = 0, 1, 2, ...$ runs $A(2^m)$ on rounds $2^m, ..., 2^{m+1} - 1$. Prove that it has a regret bound $\frac{\sqrt{2}}{\sqrt{2}-1}\alpha\sqrt{T}$.

3) Find a hypothesis class and a sequence of inputs for with the regret bound of the halving algorithm is tight.

4) Consider running FoReL with $R(w) = \frac{1}{\eta}||w||_2^2$ regularization when $w \in \mathcal{S}$ a convex set. Show that the update rule is gradient descend with lazy projections, i.e. set $\theta_1 = 0$ then $w_t = \arg\min_{w \in \mathcal{S}} ||w - \eta\theta_t||_2^2$ and $\theta_t = \theta_{t-1} - z_t$ where $z_t \in \partial f_t$.