SVM

Chcem všechny body rozdělit podrovinou, která má co největší vzdálenost od všech bodů z datasetu. Tzn. hledáme w, b takové, že

$$argmax_{w,b} \frac{1}{||w||} min_i(t_i(\phi(x_i)^T w + b)$$

Vzdálenost od přímiky můžeme přeškálovat tak, aby nejbližší bod byl vzdálený 1, celý vzorec můžeme přepsat na

$$argmin_{w,b} \frac{1}{2} ||w||^2 \text{ pokud } t_i y(x_i) \geq 1$$

KKT

Lagrangian takové funkce pak vypadá

$$L = \sum_i a_i - \sum_i \sum_j a_i a_j t_i t_j \varphi(x_i) \varphi(x_j)$$

ale derivace nám stanovuje nějaké podmínky

$$a_i \geq 0, t_i y(x_i) - 1 \geq 1, a_i(t_i y(x_i) - 1) = 0$$

To znamená, že bod je buď na rozhodovací hranici, nebo $a_i = 0$. Bodům na rozhodovací hranici můžeme říkát podpůrné vektory, $y(x) = \sum_i a_i t_i K(x, x_i) +$ b$

Lineárné neseparabilní

Kdybychom požadovali po modelu, abyc všechna data klasifikoval správně, model by mohl začít overfitovat a to nechcem. Místo toho vytvoříme soft margin, tzn. nebude nám vadit, když nějaká data budou klasifikovaná špatně, za to budeme mít penaltu ξ_i. Cheme tedy najít

$$argmin_{w,c} C \sum_i \xi_i + \frac{1}{2} ||w||^2 \text{ pokud } t_i y(x_i) \geq 1 - \xi_i, \xi_i \geq 0$$

Optimalizujeme stejnou funkci jako předtím, ale navíc máme podmíku $C \geq a_i$