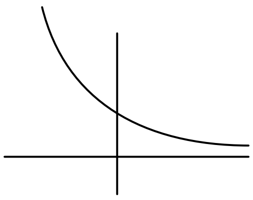
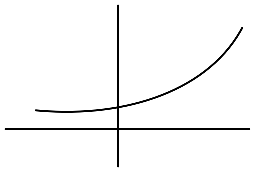
SVM

1. Logistic:

cost: $-(y\log h_\theta(x)+(1-y)\log(1-h_\theta(x)))$，分类可计算概率，



$y=1, \theta^T x \gg 0$



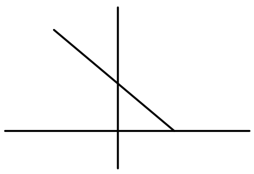
$y=0, \theta^T x \ll 0$

线性问题：两者可以互换
非线性：核函数，只能 SVM

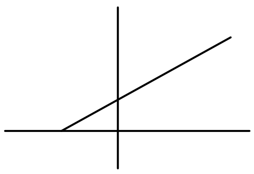
2. SVM

（1）线性可分问题

$$\min_\theta c\sum[y^{(i)}\cos t_1(\theta^T x^{(i)})+(1-y^{(i)})\cos t_0(\theta^T x^{(i)})]+\frac{1}{2}\sum\theta_i^2$$



$y=1, \theta^T x \geq 1$



$y=-1, \theta^T x \leq -1$

目标: $\min\frac{1}{2}\|w\|^2$ s.t. $y_i(wx_i+b)\geq 1$

转换为对偶问题：$\max_\alpha(\sum\alpha_i-\frac{1}{2}\sum_{i=1}^m\sum_{j=1}^m\alpha_i\alpha_j y_i y_j x_i^T x_j)$， s.t.: $\sum_{i=1}^m\alpha_i y_i=0$，$\alpha_i\geq 0$

然后使用 SMO 方法解，分为两步：

步骤一：选择α_i,α_j，α_i为违反 KKT 最大，α_j：$\left|E_i-E_j\right|$最大，

步骤二：固定其他参数，求解对偶式，更新α_i,α_j；

（2）线性不可分：　使用核函数，升维打击：

$K(x_i, x_j) =< \phi_i, \phi_j >$,原来输入高维映射的内积；

思想：在函数下的相似性，在两个维度一样，

常用核函数：

高斯核函数：$\exp(-\frac{\|x_i - x_j\|^2}{2\delta^2})$

3. 学习到的参数

$w = \sum \alpha_i y^{(i)} x^{(i)}$，对于大部分点有：$\alpha_i = 0$，不作用，对于少部分$\alpha_i > 0$，有 $y^{(i)}(w^T x + b) = 1$，为支持向量；

4. 损失函数

Hinge 损失：$\max(1 - z, 0)$，体现在，如果点远离支持向量，或者为支持向量，没有损失，如果点在间隔内，有损失；

5. Predict：

$w^T x + b = \sum \alpha_i y^{(i)} < x^{(i)}, x > + b$，其中 $x^{(i)}$ 为支持向量；