SVM的参数推导.

Step1：定义分类超平面 A：$w^Tx+b=0$，

可得法向量 w，单位法向量为 $\frac{w}{||w||}$.

边界超平面：$\begin{cases} w^Tx+b=l \\ w^Tx+b=-l \end{cases}$ ⇒ 归一化后：$\begin{cases} w^Tx+b=1 \\ w^Tx+b=-1 \end{cases}$.

设置二分类标签：$\begin{cases} y_i=1，则\ w^Tx_i+b\geq 1 \\ y_i=-1，则\ w^Tx_i+b\leq -1 \end{cases}$ ⇒ $y_i(w^Tx_i+b)\geq 1$.

Step2：边界超平面 A_1：$w^Tx+b=1$ 上任意一点x_0到 A 的距离 $d_1=\frac{w^Tx_0+b}{||w||}=\frac{1}{||w||}$.

∴ 间隔 $margin=\frac{2}{||w||}$，要间隔最大化，即求 w 的最小值，

极值问题转化：$f(w)=\frac{1}{2}||w||^2$，加入平方是为了方便计算和求模长.

∴ SVM → 最优化问题：$\begin{cases} \min\limits_{w,b} \frac{1}{2}||w||^2 \\ y_i(w^Tx_i+b)\geq 1 \end{cases}$，$i=1,2,\cdots,N$.

Step3：把最优化问题转为拉格朗日函数形态

$$L(w,b,\alpha)=\frac{1}{2}||w||^2-\sum_{i=1}^{N}\alpha_i(y_i(w^Tx_i+b)-1) \quad (\alpha_i\geq 0).$$

∴ 目标函数：$\min\limits_{w,b}\max\limits_{\alpha} L(w,b,\alpha)$，$(\alpha_i\geq 0)$

把函数第二部分当作惩罚项，则：

$\begin{cases} 若\ y_i(w^Tx_i+b)>1\ 时，L(w,b,\alpha)最大化，则\ \alpha=0. \\ 若\ y_i(w^Tx_i+b)<1\ 时，L(w,b,\alpha)最大化，则\ \alpha=+\infty，此时惩罚过大， \end{cases}$

则整体无法取到最小值.

Step4：化简 $L(w,b,\alpha)$.

$$L(w,b,\alpha)=\frac{1}{2}\|w\|^2-\sum_{i=1}^{N}\alpha_i\left(y_i\left(w^Tx_i+b\right)-1\right)$$

$$=\frac{1}{2}\|w\|^2-\sum_{i=1}^{N}\left(\alpha_iy_iw^Tx_i+\alpha_iy_ib-\alpha_i\right)$$

$$=\frac{1}{2}w^T\cdot w-\sum_{i=1}^{N}(\alpha_iy_iw^Tx_i)-\sum_{i=1}^{N}(\alpha_iy_ib)+\sum_{i=1}^{N}\alpha_i$$

∵ 要求极值，先对参数 w 和 b 求偏导.

$$\therefore\begin{cases}\dfrac{\partial L}{\partial w}=w-\sum_{i=1}^{N}\alpha_iy_ix_i=0 \Rightarrow w=\sum_{i=1}^{N}\alpha_iy_ix_i\\[2ex] \dfrac{\partial L}{\partial b}=-\sum_{i=1}^{N}\alpha_iy_i=0 \Rightarrow \sum_{i=1}^{N}\alpha_iy_i=0.\end{cases}$$

∵ 结果中带有未知参数 α，因此无法快速求解.

Step5：引入拉格朗日对偶函数.

$$L(x,\alpha)=f(x)+\sum_{i=1}^{q}\alpha_ih_i(x)$$ ，~~且对偶差异 $\Delta=\min_x L(x,\alpha)$~~

且对偶差异：$\Delta=\min\limits_{x}L(x,\alpha)-\max\limits_{\alpha}g(\alpha)$

强对偶关系需 KKT条件：

$$\begin{cases}\dfrac{\partial L}{\partial x_i}=0, & \forall i=1,2,\cdots,d\\ h_i(x)\le 0, & \forall i=1,2,\cdots,q\\ \alpha_i\ge 0, & \forall i=1,2,\cdots,q\\ \alpha_ih_i(x)=0, & \forall i=1,2,\cdots,q.\end{cases}$$

对于第四个条件：即：$\alpha_i(w^Tx_i+b)-1)=0$.

若 x_i 在边界超平面上：则 $(w^Tx_i+b)-1=0$，即假设为两个边界中的一个，此时 x_i 也叫支持向量，否则则 α_i 必须为0，

~~因此也说~~

Step6. 进行转换：

由 Step 5 可知：已满足强对偶关系，则 $\Delta=0$.

$$\Rightarrow \min_{w,b}\max_{\alpha_i\ge 0}L(w,b,\alpha)=\max_{\alpha_i\ge 0}g(\alpha)$$

由 $L(w,b,\alpha)$ 对 w,b 参数求导为0，即求最小值，代入可得：

$$L(w,b,\alpha)=\frac{1}{2}w^T\cdot w-w^T\cdot\left(\sum_{i=1}^{N}\alpha_iy_ix_i\right)-b\left(\sum_{i=1}^{N}\alpha_iy_i\right)+\sum_{i=1}^{N}\alpha_i$$

$$=\frac{1}{2}w^T\cdot w-w^T\cdot w-0+\sum_{i=1}^{N}\alpha_i,$$

$$=-\frac{1}{2}w^T\cdot w+\sum_{i=1}^{N}\alpha_i$$ ，再次代入，将 w 来自两个不同的样本点.

$$=-\frac{1}{2}\left(\sum_{i=1}^{N}\alpha_iy_ix_i^T\right)\cdot\left(\sum_{j=1}^{N}\alpha_jy_jx_j\right)+\sum_{i=1}^{N}\alpha_i$$

$$=-\frac{1}{2}\sum_{i,j=1}^{N}\alpha_iy_ix_i^T\cdot\alpha_jy_jx_j+\sum_{i=1}^{N}\alpha_i$$

再把矩阵相乘转为内积形式：∴ $L_d = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{N} \alpha_i \alpha_j y_i y_j x_i \cdot x_j$.

此时 L_d 就是只含参数 α 的关于 $L(w,b,\alpha)$ 的对偶函数 $g(\alpha)$，

且 L_d 即 $\min_{w,b} L(w,b,\alpha)$.

∴ $\min_{w,b} \max_{\alpha_i \geq 0} L(w,b,\alpha) = \max_{\alpha_i \geq 0} g(\alpha) = \max_{\alpha_i \geq 0} \min_{w,b} L(w,b,\alpha)$

∴ 目标函数变为 $\max_{\alpha_i \geq 0} \left(\sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{N} \alpha_i \alpha_j y_i y_j x_i \cdot x_j \right)$.

从而可以解出 α，$\begin{cases} 再代入 \; w = \sum_{i=1}^{N} \alpha_i y_i x_i，可求出 \; w， \\ 再代入 分类超平面 \; w^T x + b = 0，可求出 \; b. \end{cases}$

∴ 得出决策函数 $f(x\_test) = \mathrm{sign}(w^T \cdot x\_test + b)$

$$= \mathrm{sign}\left(\sum_{i=1}^{N} \alpha_i y_i x_i \cdot x\_test + b\right) .$$

∴ 对于二分类问题：$\mathrm{sign}(h) = \begin{cases} 1 & , h > 0 \\ -1 & , h < 0 \end{cases}$.