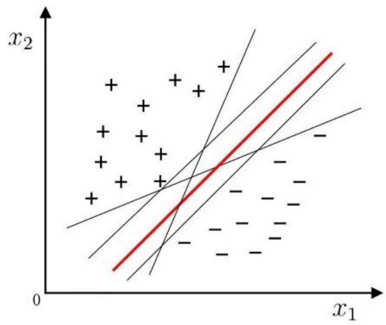
线性模型与朝平面

线性模型：在样本空间中寻找一个超平面，将不同类别的样本分开。

-Q:将训练样本分开的超平面可能有很多，哪一个好呢?



-A:应选择"正中间"，容忍性好，鲁棒性高，泛化能力最强.

点到超平面的距离（시험에 나왔었음, 공식 외우고 시간되면 증명 봐주기）:

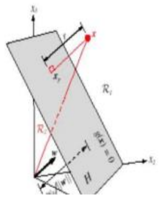
点到超平面的距离

【证明】：对于任意一点$\boldsymbol{x}_0=(x_1^0,x_2^0,\dots,x_n^0)^{\mathrm{T}}$，设其在超平面$\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}+b=0$上的投影点为$\boldsymbol{x}_1=(x_1^1,x_2^1,\dots,x_n^1)^{\mathrm{T}}$，则$\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_1+b=0$，且向量$\overrightarrow{\boldsymbol{x}_1\boldsymbol{x}_0}$与法向量$\boldsymbol{w}$平行，因此

$$|\boldsymbol{w}\cdot\overrightarrow{\boldsymbol{x}_1\boldsymbol{x}_0}|=\left|\|\boldsymbol{w}\|\cdot\cos\pi\cdot\|\overrightarrow{\boldsymbol{x}_1\boldsymbol{x}_0}\|\right|=\|\boldsymbol{w}\|\cdot\|\overrightarrow{\boldsymbol{x}_1\boldsymbol{x}_0}\|=\|\boldsymbol{w}\|\cdot r$$

$$\begin{aligned}\boldsymbol{w}\cdot\overrightarrow{\boldsymbol{x}_1\boldsymbol{x}_0}&=w_1(x_1^0-x_1^1)+w_2(x_2^0-x_2^1)+\dots+w_n(x_n^0-x_n^1)\\&=w_1x_1^0+w_2x_2^0+\dots+w_nx_n^0-(w_1x_1^1+w_2x_2^1+\dots+w_nx_n^1)\\&=\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_0-\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_1\\&=\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_0+b\end{aligned}$$

$$|\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_0+b|=\|\boldsymbol{w}\|\cdot r\Rightarrow r=\frac{|\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}+b|}{\|\boldsymbol{w}\|}$$



梯度下降法

梯度下降法：**求解无约束优化问题**的一种最常用的方法（**迭代算法**）

간략한 소개: 优化算法이고, 함수의 최소값과 최대값을 찾는데 사용됨. 특히 深度学习에서 모델훈련에 사용됨

경사 하강법은 손실 함수의 경사를 계산해 파라미터(参数)를 조정하며, 최적의 값을 점진적으로 찾아가는 방식이다.

기본 步骤:

1. 初始化：모델의 초기 파라미터를 설정
2. 计算目标函数的梯度：손실 함수의 기울기를 계산
3. 跟新参数，沿着梯度的方向移动: 계산되 기울기를 기준으로 반대 방향으로 파라미터를 조정
4. 重复2，3 ：과정을 반복하다가 기울기가 충분히 작아지거나 최대 반복 횟수에 도달하면 중지

종류: 批量梯度下降，随机梯度下降，小批量梯度下降

神经网络의 权重&偏置을 优化하는데 사용됨.

感知机模型(神经网络基础)

感知机가 무엇이냐: 是二分类的线性分类模型（判别模型）,特征空间에서 실제 예시들을 正负로 나누는 分离超平面임.

그렇다면 感知机学习는 求出将训练数据进行线性划分的分离超平面라는 뜻임. 그렇기 때문에 오분류된 손실함수를 토대로, 梯度下降法를 사용해서 손실함수를 최소화하여 感知机模型을 얻는 것

定义 2.1（感知机） 假设输入空间（特征空间）是 $\mathcal{X} \subseteq \boldsymbol{R}^n$，输出空间是 $\mathcal{Y} = \{+1, -1\}$。输入 $x \in \mathcal{X}$ 表示实例的特征向量，对应于输入空间（特征空间）的点；输出 $y \in \mathcal{Y}$ 表示实例的类别。由输入空间到输出空间的如下函数

$$f(x) = \operatorname{sign}(w \cdot x + b) \tag{2.1}$$

称为感知机。其中，w 和 b 为感知机模型参数，$w \in \boldsymbol{R}^n$ 叫作权值（weight）或权值向量（weight vector），$b \in \boldsymbol{R}$ 叫作偏置（bias），$w \cdot x$ 表示 w 和 x 的内积。sign 是符号函数，即

$$\operatorname{sign}(x) = \begin{cases} +1, & x \geqslant 0 \\ -1, & x < 0 \end{cases} \tag{2.2}$$

여기서 w*x + b가 **超平面**에 해당하고, w가 超平面의 **法向量**, b가 超平面的**截距**에 해당함

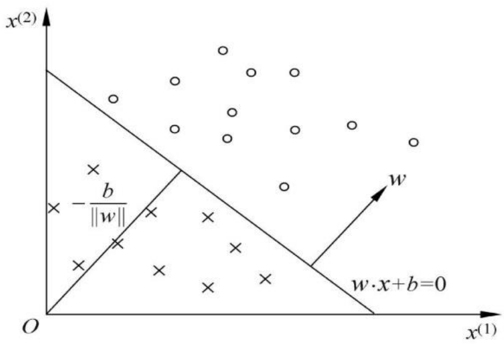


图 2.1　感知机模型

线性可分数据集: 쉽게 말해서 데이터 셋을 완벽하게 正实例点이랑 负实例点으로 나눌 수 있는 超平面이 있으면 그 데이터 셋은 线性可分한거임

그리고 데이터가 线性可分이라고 하면 상응하는 超平面을 찾는게 感知机学习의 목표임. 그리고 그걸 찾기 위해 즉 w와 b를 확정하기 위해 손실함수를 정의하고, 이 손실함수를 极小化하는 거임

그 손실함수:

$$L(w,b) = -\sum_{x_i \in M} y_i (w \cdot x_i + b)$$

感知机学习算法 = 随机梯度下降法임의로 超平面(w0, b0)을 골라서 梯度下降法로 极小化目标函数

算法的收敛性: 감자기 알고리즘음 初值이나 误分类点을 다르게 고르기 때문에 답이 다를수있음 근데 经过有限次数迭代可以得到一个将数据集完全正确划分的分离超平面及感知机模型。

유일한 超平面을 얻고 싶으면 제약조건을 걸어야함 → 支持向量机의 사상