19 From Lines to Hyperplanes 直线到超平面

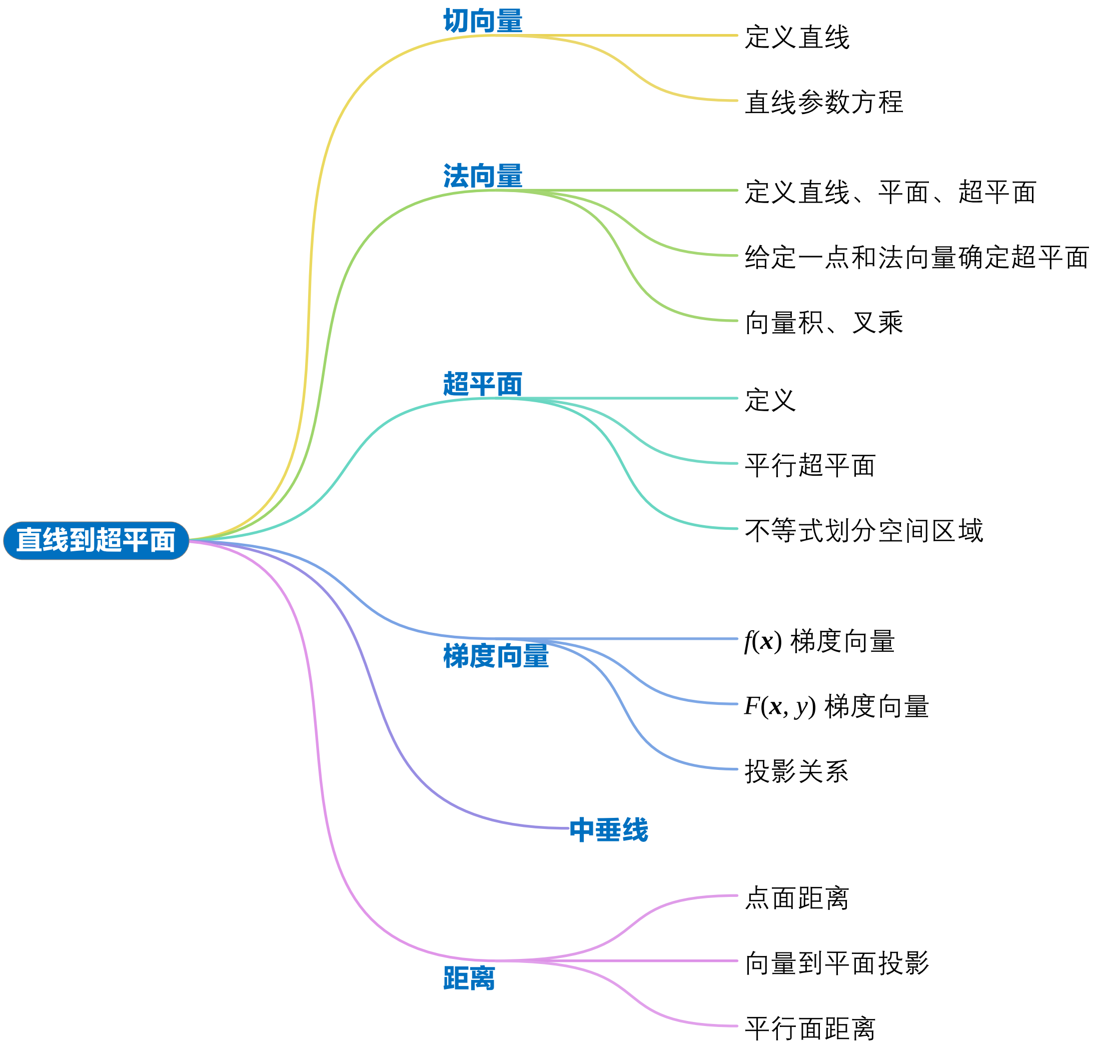
用线性代数工具分析直线、平面和超平面

古人说，算数和几何是数学的双翼；而我认为，算数和几何是任何量化科学的基础和精髓。不仅如此，它们还是压顶石。任何科学的结果都需要用数字或者几何图形来表达。将结果转化为数字，需要借助算数；将结果转化为图形，需要借助几何。

An ancient writer said that arithmetic and geometry are the wings of mathematics; I believe one can say without speaking metaphorically that these two sciences are the foundation and essence of all the sciences which deal with quantity. Not only are they the foundation, they are also, as it were, the capstones; for, whenever a result has been arrived at, in order to use that result, it is necessary to translate it into numbers or into lines; to translate it into numbers requires the aid of arithmetic, to translate it into lines necessitates the use of geometry.

—— 约瑟夫·拉格朗日 (Joseph Lagrange) | 法国籍意大利裔数学家和天文学家 | 1736 ~ 1813

- `matplotlib.pyplot.contour()` 绘制等高线图
- `matplotlib.pyplot.quiver()` 绘制箭头图
- `numpy.meshgrid()` 产生网格化数据
- `numpy.ones_like()` 用来生成和输入矩阵形状相同的全 1 矩阵
- `subs()` 符号代数式中替换
- `sympy.abc import x` 定义符号变量 x
- `sympy.diff()` 求解符号导数和偏导解析式
- `sympy.evalf()` 将符号解析式中未知量替换为具体数值
- `sympy.lambdify()` 将符号表达式转化为函数
- `sympy.plot_implicit()`绘制隐函数方程
- `sympy.simplify()` 简化代数式
- `sympy.symbols()` 定义符号变量


直线到超平面
切向量
定义直线
直线参数方程
法向量
定义直线、平面、超平面
给定一点和法向量确定超平面
向量积、叉乘
超平面
定义
平行超平面
不等式划分空间区域
梯度向量
f(x) 梯度向量
F(x, y) 梯度向量
投影关系
中垂线
距离
点面距离
向量到平面投影
平行面距离

19.1 线性代数视角看几何

至此，我们已经掌握大量线性代数运算工具。向量天然具备几何属性，这使得线性代数和几何之间的联系显而易见。本书前文利用几何视角帮助我们可视化大量重要的线性代数概念，让众多枯燥的概念和运算变得栩栩如生。

包括本章在内的接下来三章则利用线性代数工具讲解数据科学、机器学习中常见的几何知识。

本系列丛书《数学要素》一册介绍大量的平面和立体几何知识，而线性代数工具可以将这些知识从二维、三维，延伸到更高维度，比如将直线的概念延伸到超平面。

19.2 切向量：可以用来定义直线

如图 1 (a) 所示，直线上任意一点**切向量** (tangent vector) 和直线重合。

图 1 (b) 中，曲线在一点处的切向量是曲线该点处切线方向上的向量。

如图 1 (c) 所示，三维空间平面上某点切线有无数条，它们都在同一个平面内。

同样，如图 1 (d) 所示，光滑曲面某点切线有无数条，它们都在曲面上该点切平面内。也可以说，这些切线构造该切平面。

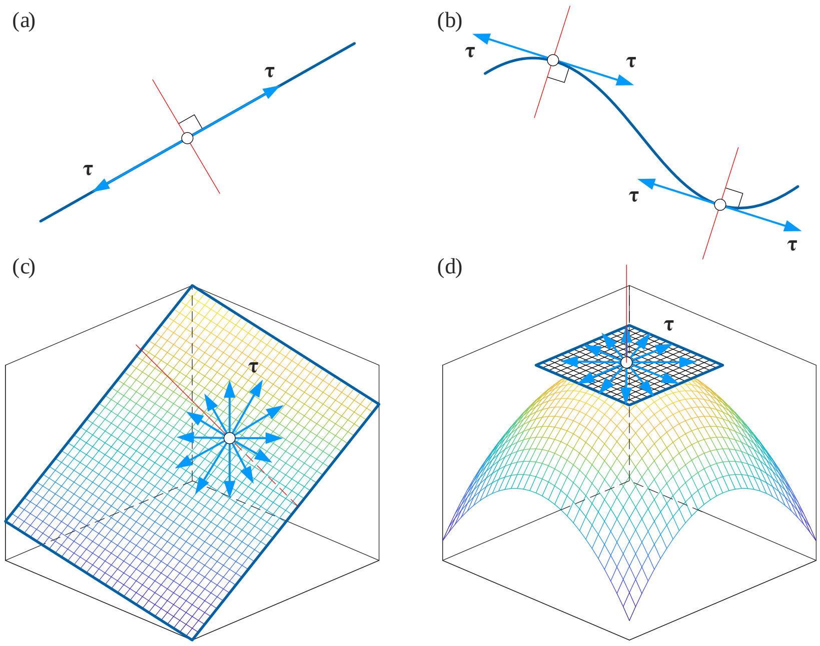


图 1. 直线、平面和光滑曲面切向量

丛书一般用 $\boldsymbol{\tau}$ 来表达切向量。**单位切向量** (unit tangent vector) $\hat{\boldsymbol{\tau}}$ 通过向量单位化获得:

$$\hat{\boldsymbol{\tau}} = \frac{\boldsymbol{\tau}}{\|\boldsymbol{\tau}\|} \tag{1}$$

单位切向量 $\hat{\boldsymbol{\tau}}$ 模为 1。

描述平面直线

切向量可以用来描述直线。给定空间一点 $\boldsymbol{c}$ 和直线的切向量 $\boldsymbol{\tau}$ 便可以确定一条直线:

$$\boldsymbol{x} = k\boldsymbol{\tau} + \boldsymbol{c} \tag{2}$$

上式实际上是前文介绍的平移。

举个例子,用切向量描述平面上直线:

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = k\begin{bmatrix} \tau_1 \\ \tau_2 \end{bmatrix} + \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \tag{3}$$

如图 1 (a) 所示,这条穿越原点、切向量为 $\boldsymbol{\tau} = [4, 3]^{\mathrm{T}}$ 的直线可以写作:

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = k\begin{bmatrix} 4 \\ 3 \end{bmatrix} \tag{4}$$

其中,k 任意实数。

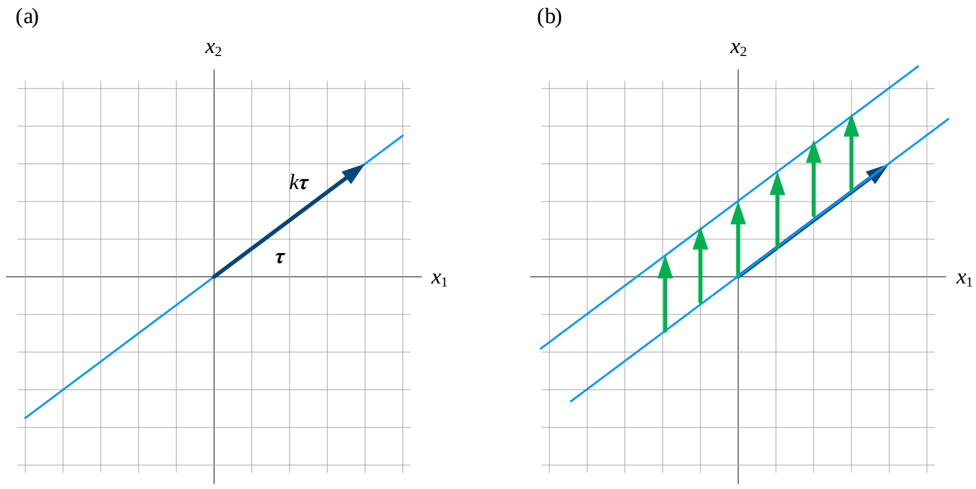


图 2. 用切向量定义平面直线

图 1 (b) 所示,(4) 直线向上平移 $[0, 2]^{\mathrm{T}}$,得到如下直线:

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = k\begin{bmatrix} 4 \\ 3 \end{bmatrix} + \begin{bmatrix} 0 \\ 2 \end{bmatrix} \tag{5}$$

将 (5) 展开得到平面直线的参数方程：

$$\begin{cases} x_1 = 4k \\ x_2 = 3k + 2 \end{cases} \tag{6}$$

用 (3) 这种方式定义平面直线的好处是，切向量可以指向任意方向，比如水平方向 $[2, 0]^{\mathrm{T}}$、竖直方向 $[0, -1]^{\mathrm{T}}$。

描述三维空间直线

类似地，如图 3 所示，给定切向量和直线通过的一点 $\boldsymbol{c}$，便可以定义一条三维空间直线：

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = k\begin{bmatrix} \tau_1 \\ \tau_1 \\ \tau_3 \end{bmatrix} + \begin{bmatrix} c_1 \\ c_3 \\ c_3 \end{bmatrix} \tag{7}$$

将 (7) 展开便得到三维空间直线的参数方程：

$$\begin{cases} x_1 = k\tau_1 + c_1 \\ x_2 = k\tau_2 + c_2 \\ x_3 = k\tau_3 + c_3 \end{cases} \tag{8}$$

上述直线定义也可以推广到高维。

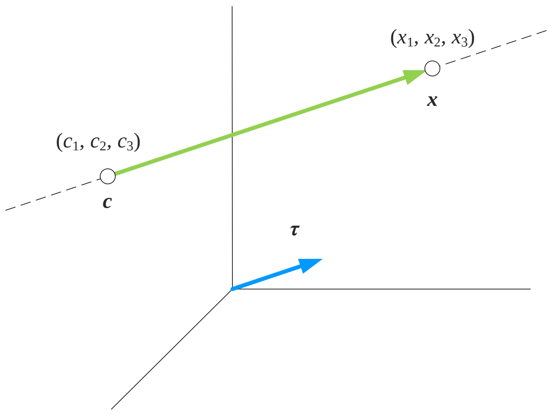


图 3. 空间直线定义

19.3 法向量：定义直线、平面、超平面

本系列丛书常用法向量定义直线、平面，甚至**超平面** (hyperplane)。直线**法向量** (normal vector) 为垂直于直线非零向量，如图 4 (a) 所示。

如图 4 (b) 所示，光滑曲线某点法向量垂直于曲线该点切线。如图 4 (c) 所示，**平面法向量** (a normal line to a surface) 垂直于平面内任意直线。光滑连续曲面某点法向量为曲面该点处**切平面** (tangent plane) 的法向量，如图 4 (d) 所示。

本章用 $\boldsymbol{n}$ 或 $\boldsymbol{w}$ 表达法向量。非零法向量 $\boldsymbol{n}$ 的**单位法向量** (unit normal vector) $\hat{\boldsymbol{n}}$ 通过单位化获得：

$$\hat{\boldsymbol{n}} = \frac{\boldsymbol{n}}{\|\boldsymbol{n}\|} \tag{9}$$

同样，单位法向量 $\hat{\boldsymbol{n}}$ 模为 1。

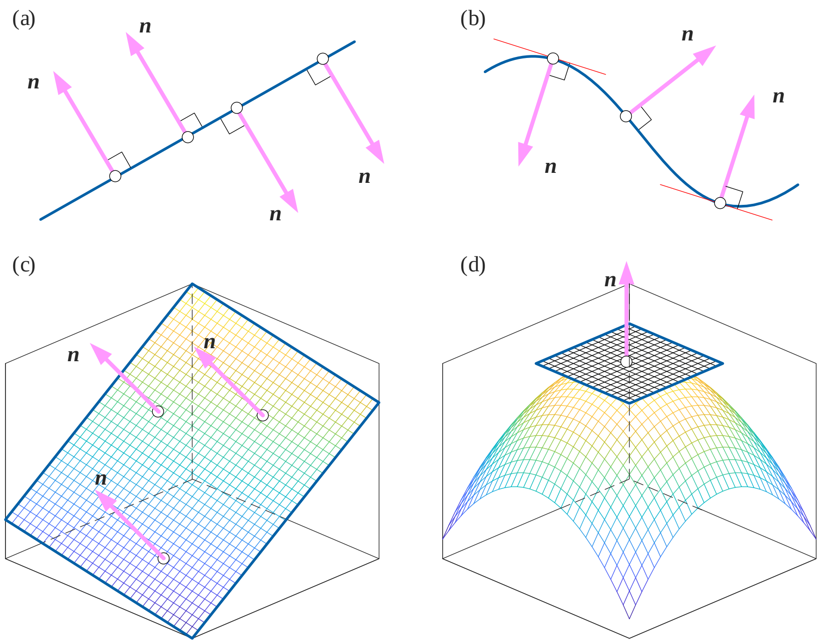


图 4. 直线、平面和光滑曲面法向量

描述三维空间平面

如图 5 所示，过空间一点与已知直线相垂直平面唯一。从向量视角来看，给定平面上一点和平面法向量 $\boldsymbol{n}$，可以确定一个平面。

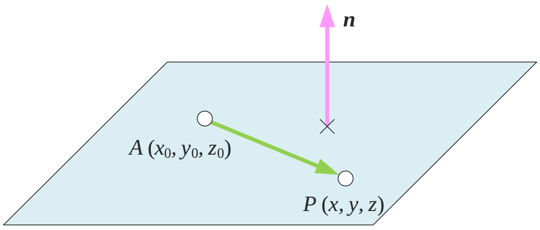


图 5. 空间平面定义

举个例子，三维空间内某个平面通过点 A (1, 2, 3) 且垂直于法向量 $\boldsymbol{n}$ = [3, 2, 1]$^{\mathrm{T}}$。

为了确定该直线解析式，定义平面上任意一点 P (x_1, x_2, x_3)，点 A 和 P 确定的向量垂直于法向量 $\boldsymbol{n}$，所以下式成立：

$$\boldsymbol{n} \cdot \overrightarrow{AP} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 - 1 \\ x_2 - 2 \\ x_3 - 3 \end{bmatrix} = 0 \tag{10}$$

整理 (10) 得到平面的解析式：

$$3x_1 + 2x_2 + x_3 - 10 = 0 \tag{11}$$

再举个例子，求通过三个点 P_1 (3, 1, 2)，P_2 (1, 2, 3)，P_3 (4, −1, 1) 的平面解析式。

$\boldsymbol{a}$ 是起点为 P_1 终点为 P_2 的向量，$\boldsymbol{b}$ 是起点为 P_1 终点为 P_3 的向量。$\boldsymbol{a}$ 和 $\boldsymbol{b}$ 分别为：

$$\boldsymbol{a} = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix} \tag{12}$$

向量 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 的向量积，即叉乘 $\boldsymbol{a} \times \boldsymbol{b}$ 结果如下：

$$\boldsymbol{a} \times \boldsymbol{b} = \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix} \tag{13}$$

如图 6 所示，$\boldsymbol{a} \times \boldsymbol{b}$ 便是平面法向量 $\boldsymbol{n}$。

有了法向量 $\boldsymbol{n}$，仅仅需要平面任意一点便可以确定平面解析式。利用 P_1 和法向量 $\boldsymbol{n}$ 可以得到如下平面解析式：

$$x_1 - x_2 + 3x_3 - 8 = 0 \tag{14}$$

P_1 (3, 1, 2)，P_2 (1, 2, 3)，P_3 (4, −1, 1) 三点都在 (14) 平面上，请大家自行验证。

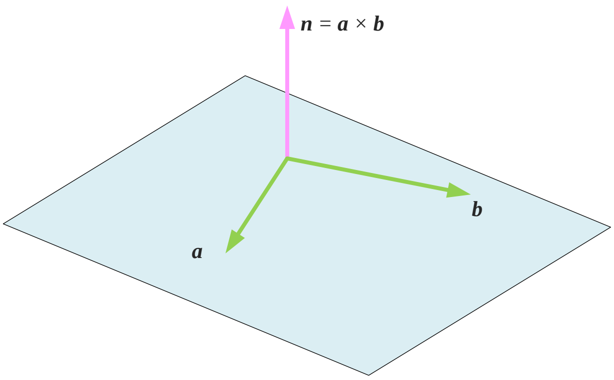


图 6. 向量叉乘为平面法向量

19.4 超平面：一维直线和二维平面的推广

本节将上一节平面扩展到多维空间。

D 维超平面的定义如下：

$$\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} + b = 0 \tag{15}$$

其中，

$$\boldsymbol{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_D \end{bmatrix}, \quad \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_D \end{bmatrix} \tag{16}$$

$\boldsymbol{w}$ 为超平面法向量，形式为列向量。(15) 中，$\boldsymbol{w}$ 和 $\boldsymbol{x}$ 长度均为 D。

(15) 也可以通过内积方式表达：

$$\boldsymbol{w} \cdot \boldsymbol{x} + b = 0 \tag{17}$$

展开 (15) 得到：

$$w_1 x_1 + w_2 x_2 + \ldots + w_D x_D + b = 0 \tag{18}$$

D = 2

特别地，$D = 2$ 时，(17) 对应的平面直线解析式为：

$$w_1 x_1 + w_2 x_2 + b = 0 \tag{19}$$

(19) 不止表达类似一次函数的直线。$w_1 = 0$ 时，(19) 表达平行于横轴的直线，类似于常数函数直线。$w_2 = 0$，(19) 为垂直横轴直线，这显然不是函数。

$D = 3$

$D = 3$ 时，(17) 对应的三维空间平面为：

$$w_1x_1 + w_2x_2 + w_3x_3 + b = 0 \tag{20}$$

图 7 所示为上述几种几何图形。平面上，法向量 $\boldsymbol{w}$ 垂直于直线。空间中，法向量 $\boldsymbol{w}$ 垂直于平面或**超平面** (hyperplane)。

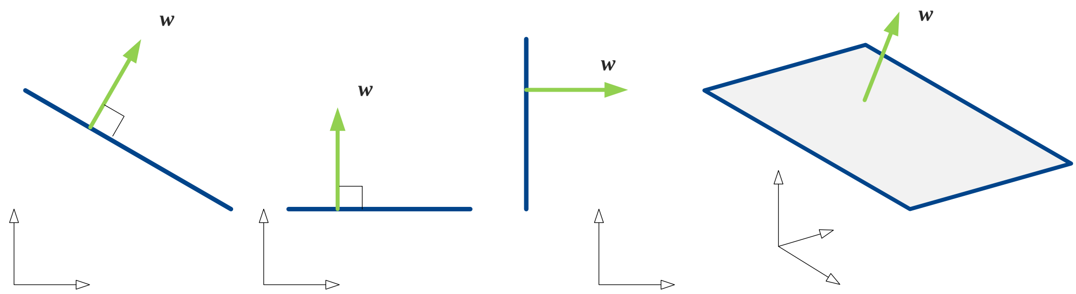


图 7. 几种特殊形态的超平面

超平面关系

如果两个超平面平行，则法向量平行或重合。如果两个超平面垂直，则法向量垂直，即内积为 0。

(19) 中 b 取不同值时，代表一系列平行直线，如图 8 (a) 所示。

而 (20) 中 b 取不同值时，代表一系列平行平面，如图 8 (b) 所示。

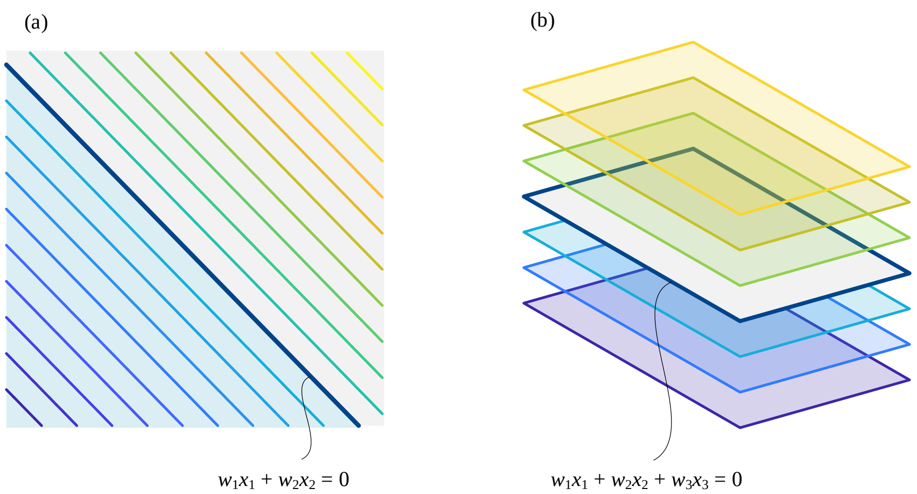


图 8. 平行直线和平行平面

划定区域

此外，某个确定的超平面解析式 $\mathbf{w}^{\mathrm{T}}\mathbf{x} + b = 0$ 可以划分空间区域。这一点在机器学习很多算法中非常重要。

图 8 (a) 中，$w_1x_1 + w_2x_2 = 0$ 将平面划分为 $w_1x_1 + w_2x_2 > 0$ 和 $w_1x_1 + w_2x_2 < 0$ 两个区域。

图 8 (b) 中，$w_1x_1 + w_2x_2 + w_3x_3 = 0$ 将空间划分为 $w_1x_1 + w_2x_2 + w_3x_3 > 0$ 和 $w_1x_1 + w_2x_2 + w_3x_3 < 0$ 两个区域。

定义多元一次函数：

$$f(\mathbf{x}) = \mathbf{w}^{\mathrm{T}}\mathbf{x} + b \tag{21}$$

超平面"上方"的数据点满足：

$$f(\mathbf{x}) = \mathbf{w}^{\mathrm{T}}\mathbf{x} + b > 0 \tag{22}$$

展开 (22) 得到：

$$w_1x_1 + w_2x_2 + \ldots + w_Dx_D + b > 0 \tag{23}$$

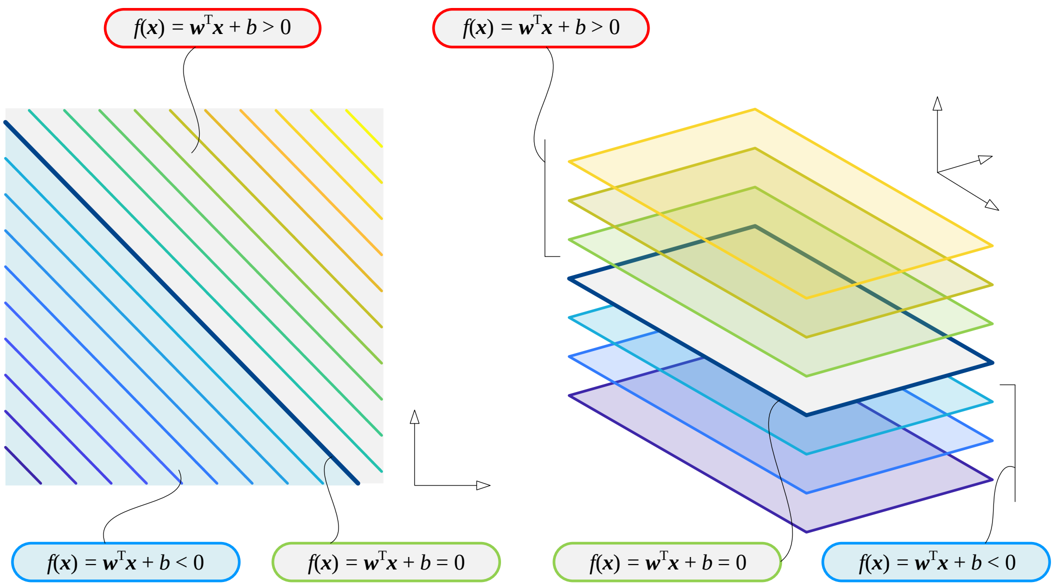


图 9. 超平面分割空间

超平面"下方"的数据点满足：

$$f(\mathbf{x}) = \mathbf{w}^{\mathrm{T}}\mathbf{x} + b < 0 \tag{24}$$

展开 (24) 得到：

$$w_1x_1 + w_2x_2 + \ldots + w_Dx_D + b < 0 \tag{25}$$

请读者注意，这里所说的“上方”和“下方”仅仅是方便大家理解。更准确地说，以 (15) 中 $f(\boldsymbol{x}) = 0$ 为基准，“上方”对应 $f(\boldsymbol{x}) > 0$，“下方”对应 $f(\boldsymbol{x}) < 0$。

在机器学习中，类似图 9 中起到划分空间作用的超平面，常常被称作决策平面。

19.5 平面与梯度向量

本节将超平面和函数联系在一起。

构造多元一次函数：

$$f(\boldsymbol{x}) = \boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} + b \tag{26}$$

$f(\boldsymbol{x}) = 0$ 对应的便是 (15) 所示超平面解析式。$f(\boldsymbol{x}) = c$ 时，相当于 (15) 所示超平面平行移动。

$f(\boldsymbol{x})$ 函数的梯度向量：

$$\nabla f(\boldsymbol{x}) = \begin{bmatrix} \dfrac{\partial f}{\partial x_1} \\ \dfrac{\partial f}{\partial x_2} \\ \vdots \\ \dfrac{\partial f}{\partial x_D} \end{bmatrix} = \boldsymbol{w} \tag{27}$$

相信大家已经发现 $f(\boldsymbol{x})$ 函数的梯度向量 $\boldsymbol{w}$ 便是本节前文的法向量。

构造新函数

令 $y = f(\boldsymbol{x})$，构造 $D + 1$ 维多元函数 $F(\boldsymbol{x}, y)$：

$$F(\boldsymbol{x}, y) = \boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} + b - y \tag{28}$$

$F(\boldsymbol{x}, y) = 0$ 相当于降维，得到 (26)。

$F(\boldsymbol{x}, y)$ 函数的梯度向量为：

$$\nabla F(\boldsymbol{x}, y) = \begin{bmatrix} \dfrac{\partial F}{\partial x_1} \\ \dfrac{\partial F}{\partial x_2} \\ \vdots \\ \dfrac{\partial F}{\partial x_D} \\ \dfrac{\partial F}{\partial y} \end{bmatrix} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_D \\ -1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{w} \\ -1 \end{bmatrix} \tag{29}$$

容易发现，(29) 和 (27) 梯度向量之间存在某种投影关系：

$$\nabla f(\boldsymbol{x}) = \begin{bmatrix} \boldsymbol{I}_{D\times D} & \boldsymbol{0}_{D\times 1} \end{bmatrix}_{D\times(D+1)} \nabla F(\boldsymbol{x}, y) \tag{30}$$

展开上式得到：

$$\nabla f(\boldsymbol{x}) = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{w} \\ -1 \end{bmatrix} = \boldsymbol{w}_{D\times 1} \tag{31}$$

上式相当于从 $D+1$ 维空间降维到 D 维空间。而 $\nabla F(\boldsymbol{x}, y)$ 就是 $F(\boldsymbol{x}, y) = 0$ 在 $D+1$ 维空间的法向量。图 10 所示为某个三维空间平面法向量 $\boldsymbol{n} = \nabla F(\boldsymbol{x}, y)$ 和梯度向量 $\nabla f(\boldsymbol{x})$ 的关系。

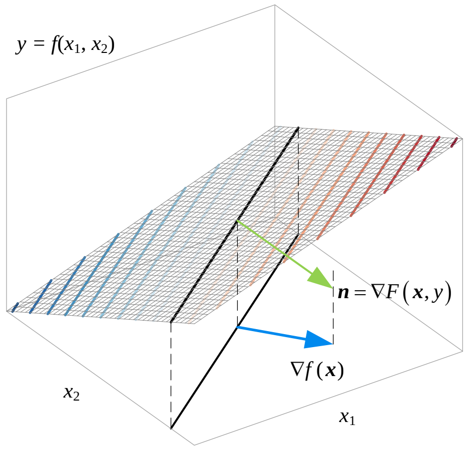


图 10. 平面法向量和梯度向量的关系

四个例子

上述投影关系对于理解很多机器学习算法至关重要，下面我用几个三维平面展开讲解上述关系。

图 11 (a) 展示的平面垂直于 x_1y 平面，具体解析式如下：

$$f(x_1, x_2) = x_1 \tag{32}$$

二元函数 $f(x_1, x_2)$ 梯度向量如下：

$$\nabla f(\boldsymbol{x}) = \begin{bmatrix} \dfrac{\partial f(\boldsymbol{x})}{\partial x_1} \\ \dfrac{\partial f(\boldsymbol{x})}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{33}$$

如图 11 (b) 所示，发现梯度向量平行于 x_1 轴，方向为 x_1 正方向，向量方向和大小不随位置变化。沿着梯度方向运动，$f(x_1, x_2)$ 不断增大。$f(x_1, x_2)$ 等高线相互平行，梯度向量和函数等高线垂直。

构造三元函数 $F(x_1, x_2, y)$：

$$F\left(x_1, x_2, y\right) = x_1 - y \tag{34}$$

$F(x_1, x_2, y)$ 梯度向量：

$$\nabla F\left(\boldsymbol{x}, y\right) = \begin{bmatrix} \dfrac{\partial F}{\partial x_1} \\ \dfrac{\partial F}{\partial x_2} \\ \dfrac{\partial F}{\partial y} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \tag{35}$$

$\nabla F(\boldsymbol{x}, y)$ 是图 11 (a) 三维平面的法向量。$\nabla F(\boldsymbol{x}, y)$ 向 x_1x_2 平面投影得到 $\nabla f(\boldsymbol{x})$，即：

$$\nabla f\left(\boldsymbol{x}\right) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \nabla F\left(\boldsymbol{x}, y\right) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{36}$$

图 11 (b) 等高线则对应一系列垂直于横轴的直线，它们可以写成：

$$x_1 + b = 0 \tag{37}$$

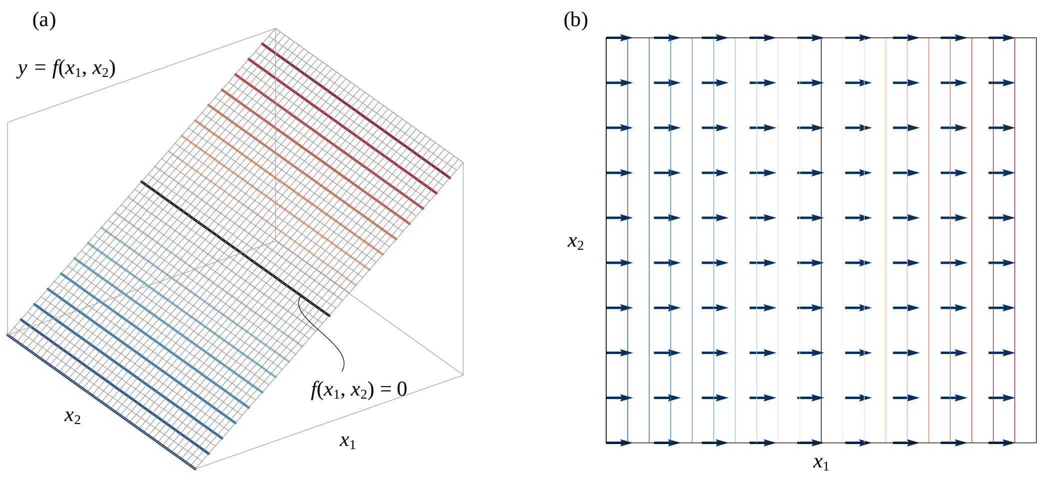


图 11. 垂直于 x_1-y 平面，梯度向量为 x_1 正方向

再举个例子，图 12 (a) 对应的二元函数 $f(x_1, x_2)$ 解析式为：

$$f\left(x_1, x_2\right) = -x_1 \tag{38}$$

$f(x_1, x_2)$ 梯度向量如下：

$$\nabla f\left(\boldsymbol{x}\right) = \begin{bmatrix} \dfrac{\partial f\left(\boldsymbol{x}\right)}{\partial x_1} \\ \dfrac{\partial f\left(\boldsymbol{x}\right)}{\partial x_2} \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix} \tag{39}$$

图 12 (b) 告诉我们，$f(x_1, x_2)$ 梯度向量同样平行于 x_1 轴，方向为 x_1 负方向，向量方向和大小不随位置变化。类似 (34)，请大家自行构造三元函数 $F(x_1, x_2, y)$，并计算它的梯度向量 $\nabla F(\boldsymbol{x}, y)$。并且分析 $\nabla F(\boldsymbol{x}, y)$ 和 (39) 的关系。

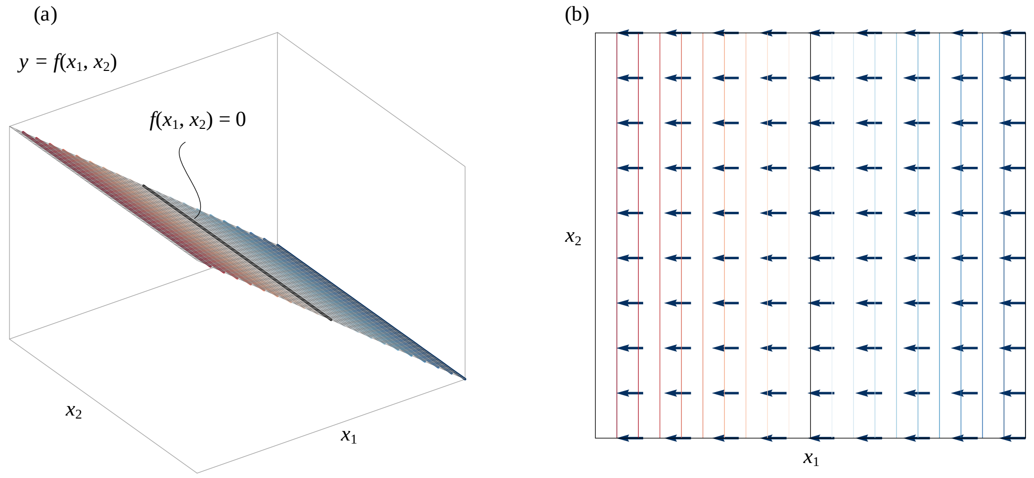


图 12. 垂直于 x_1y 平面，梯度向量为 x_1 负方向

图 13 展示平面解析式 $f(x_1, x_2)$ 如下：

$$f\left(x_1, x_2\right) = x_2 \tag{40}$$

$f(x_1, x_2)$ 梯度向量如下：

$$\nabla f\left(\boldsymbol{x}\right) = \begin{bmatrix} \dfrac{\partial f\left(\boldsymbol{x}\right)}{\partial x_1} \\ \dfrac{\partial f\left(\boldsymbol{x}\right)}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{41}$$

如图 13 (b) 所示，$f(x_1, x_2)$ 梯度向量平行于 x_2 轴，方向为 x_2 正方向。也请大家构造其三元函数 $F(x_1, x_2, y)$，同时计算它的梯度向量 $\nabla F(\boldsymbol{x}, y)$。

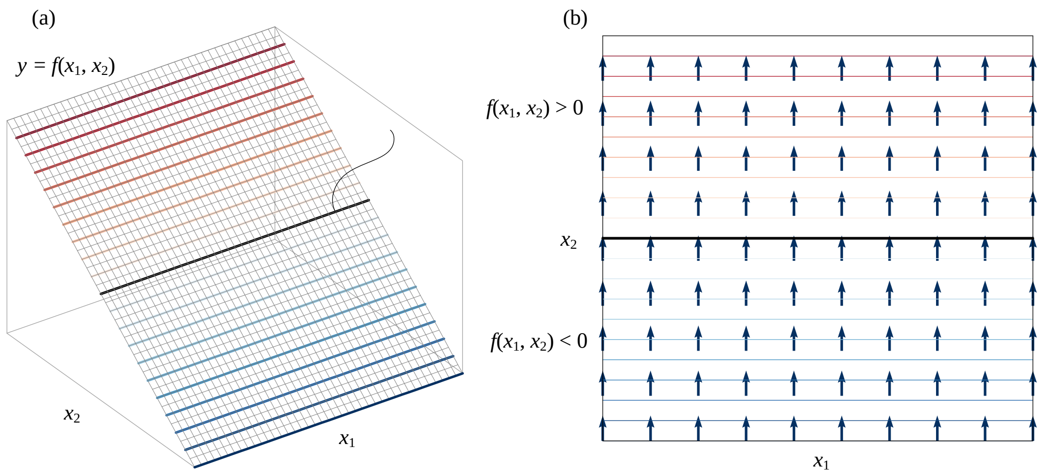


图 13. 垂直于 x_2y 平面，梯度向量为 x_2 正方向

最后一个例子，图 14 (a) 平面解析式如下：

$$f(x_1, x_2) = x_1 + x_2 \tag{42}$$

$f(x_1, x_2)$ 梯度也是一个固定向量，具体如下：

$$\nabla f(\boldsymbol{x}) = \begin{bmatrix} \dfrac{\partial f(\boldsymbol{x})}{\partial x_1} \\ \dfrac{\partial f(\boldsymbol{x})}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{43}$$

如图 14 所示，梯度向量和 x_1 轴正方向夹角为 45°，指向右上方。沿着此梯度方向运动，$f(x_1, x_2)$ 不断增大。请大家按照上述思路分析图 14 平面。

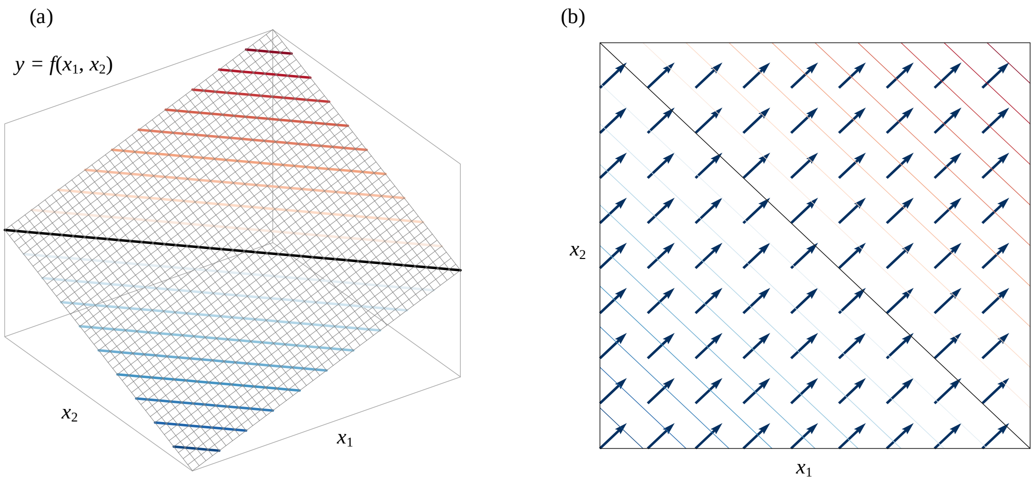


图 14. $f(x_1, x_2) = x_1 + x_2$ 平面和梯度

请读者自行修改 Bk4_Ch19_01.py，并绘制图 11 ~ 图 14 几幅图像。

```
# Bk4_Ch19_01.py

import sympy

#define symbolic vars, function
x1,x2 = sympy.symbols('x1 x2')
# f_x = 3*x1**2-5*x2**2
f_x = x1 - x2

#take the gradient symbolically
grad_f = [sympy.diff(f_x,var) for var in (x1,x2)]

f_x_fcn = sympy.lambdify([x1,x2],f_x)

#turn into a bivariate lambda for numpy
grad_fcn = sympy.lambdify([x1,x2],grad_f)
```

```
import numpy as np
import matplotlib.pyplot as plt

xx1, xx2 = np.meshgrid(np.linspace(-4,4,41),np.linspace(-4,4,41))

# coarse mesh
xx1_, xx2_ = np.meshgrid(np.linspace(-4,4,10),np.linspace(-4,4,10))
V = grad_fcn(xx1_,xx2_)
V_z = np.ones_like(V[1]);

ff_x = f_x_fcn(xx1,xx2)
# ff_x_ = f_x_fcn(xx1_,xx2_)

color_array = np.sqrt(V[0]**2 + V[1]**2)
l_3D_vectors = np.sqrt(V[0]**2 + V[1]**2 + V_z**2)

# 3D
ax = plt.figure().add_subplot(projection='3d')
ax.plot_wireframe(xx1, xx2, ff_x, rstride=1,
                  cstride=1, color = [0.5,0.5,0.5],
                  linewidth = 0.2)
ax.contour3D(xx1, xx2, ff_x,20)
ax.contour3D(xx1, xx2, ff_x, levels = 0, colors = 'k')

# plt.quiver (xx1_, xx2_, ff_x_, V[0], V[1], V_z,length = .5)
ax.xaxis.set_ticks([])
ax.yaxis.set_ticks([])
ax.zaxis.set_ticks([])
plt.xlim(-4,4)
plt.ylim(-4,4)
ax.view_init(30, -125)
ax.set_xlabel('$x_1$')
ax.set_ylabel('$x_2$')
ax.set_zlabel('$f(x_1,x_2)$')
plt.tight_layout()


# 2D
fig, ax = plt.subplots()
plt.quiver (xx1_, xx2_, V[0], V[1], color_array,
            angles='xy', scale_units='xy',scale=2,
            edgecolor='none', facecolor= 'b')

plt.contour(xx1, xx2, ff_x,20)
plt.contour(xx1, xx2, ff_x, levels = 0, colors = 'k')
plt.show()
ax.set_aspect('equal')
ax.xaxis.set_ticks([])
ax.yaxis.set_ticks([])
ax.set_xlabel('$x_1$')
ax.set_ylabel('$x_2$')
plt.tight_layout()
```

19.6 中垂线：用向量求解析式

本系列丛书《数学要素》介绍过中垂线。两点构成一条线段，中垂线 (perpendicular bisector) 通过该线段中点，且垂直线段。

本节介绍如何利用向量求解中垂线解析式。

如图 15 所示，$\boldsymbol{x}$ 代表中垂线上任意一点，中垂线通过 $\boldsymbol{\mu}_1$ 和 $\boldsymbol{\mu}_2$ 中点 $(\boldsymbol{\mu}_2 + \boldsymbol{\mu}_1)/2$。

令 $\boldsymbol{x}$ 和中点$(\boldsymbol{\mu}_2+\boldsymbol{\mu}_1)/2$ 构成的向量为 $\boldsymbol{a}$:

$$\boldsymbol{a}=\boldsymbol{x}-\frac{1}{2}(\boldsymbol{\mu}_2+\boldsymbol{\mu}_1) \tag{44}$$

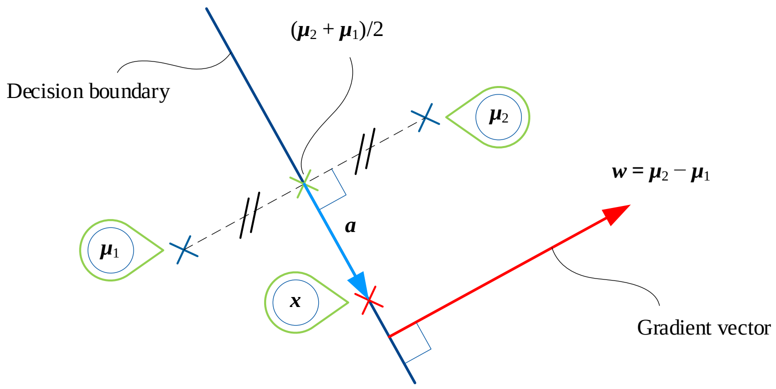


图 15. $\boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2$时，中垂线位置

$(\boldsymbol{\mu}_2-\boldsymbol{\mu}_1)$ 为中垂线法向量，它垂直 $\boldsymbol{a}$，所以下式成立:

$$(\boldsymbol{\mu}_2-\boldsymbol{\mu}_1)\cdot\boldsymbol{a}=(\boldsymbol{\mu}_2-\boldsymbol{\mu}_1)^{\mathrm{T}}\boldsymbol{a}=0 \tag{45}$$

将 (44) 代入 (45)，得到:

$$(\boldsymbol{\mu}_2-\boldsymbol{\mu}_1)\cdot\left[\boldsymbol{x}-\frac{1}{2}(\boldsymbol{\mu}_2+\boldsymbol{\mu}_1)\right]=0 \tag{46}$$

展开得到中垂线解析式:

$$\underbrace{(\boldsymbol{\mu}_2-\boldsymbol{\mu}_1)^{\mathrm{T}}}_{\text{Norm vector}}\boldsymbol{x}\underbrace{-\frac{1}{2}(\boldsymbol{\mu}_2-\boldsymbol{\mu}_1)^{\mathrm{T}}(\boldsymbol{\mu}_2+\boldsymbol{\mu}_1)}_{\text{Constant}}=0 \tag{47}$$

举个例子

平面上一条直线为 (1, 2) 和 (3, 4) 两点的中垂线，容易知道这条直线的法向量为:

$$\boldsymbol{w}=\begin{bmatrix}3\\4\end{bmatrix}-\begin{bmatrix}1\\2\end{bmatrix}=\begin{bmatrix}2\\2\end{bmatrix} \tag{48}$$

中垂线通过(1, 2) 和 (3, 4) 两点的中点 (2, 3)。这样有了法向量和直线上一点，就可以构造如下等式:

$$\begin{bmatrix}2\\2\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}x_1-2\\x_2-3\end{bmatrix}=0 \tag{49}$$

整理得到中垂线的解析式:

$$x_1 + x_2 - 5 = 0 \tag{50}$$

19.7 用向量计算距离

本节要介绍两个重要距离——点面距离，平行面距离。本节内容对于理解很多机器学习算法特别重要，请大家务必认真对待。建议大家跟着本节思路一起推导公式。

点面距离

图 16 所示，直线、平面或超平面上任一点为 $\boldsymbol{x}$，满足下式：

$$\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} + b = 0 \tag{51}$$

下面讲解如何用线性代数工具计算图 16 中一点 $\boldsymbol{q}$ 到 (51) 距离。

整理 (51) 得到：

$$\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} = -b \tag{52}$$

直线、平面或超平面上取任意一点 $\boldsymbol{x}$，$\boldsymbol{x}$ 和 $\boldsymbol{q}$ 构造的向量为 $\boldsymbol{a}$：

$$\boldsymbol{a} = \boldsymbol{q} - \boldsymbol{x} \tag{53}$$

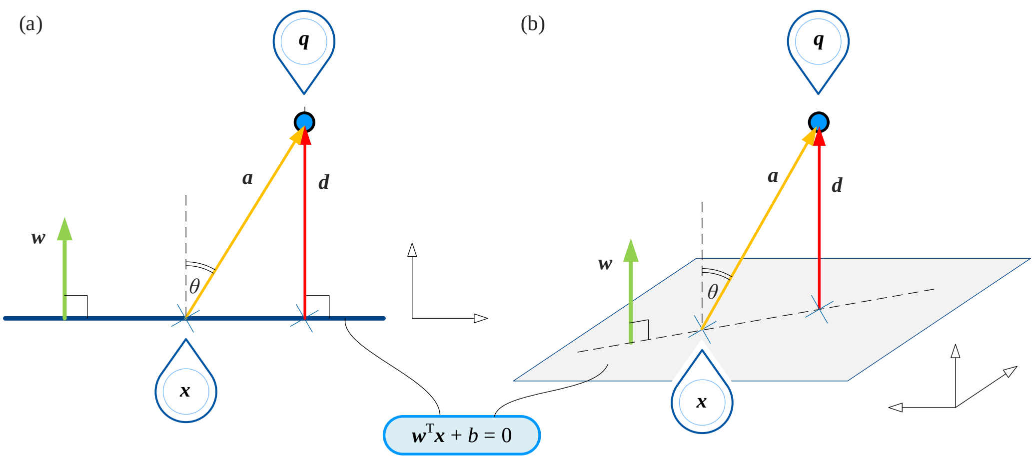


图 16. 直线外一点到直线距离，和平面外一点到平面距离

向量 $\boldsymbol{a}$ 向梯度向量 $\boldsymbol{w}$ 方向投影，可以得到向量 $\boldsymbol{d}$：

$$\boldsymbol{d} = \|\boldsymbol{a}\|\cos\theta \frac{\boldsymbol{w}}{\|\boldsymbol{w}\|} = \|\boldsymbol{a}\| \frac{\boldsymbol{w}^{\mathrm{T}}\boldsymbol{a}}{\|\boldsymbol{a}\|\|\boldsymbol{w}\|} \frac{\boldsymbol{w}}{\|\boldsymbol{w}\|} = \frac{\boldsymbol{w}^{\mathrm{T}}\boldsymbol{a}}{\|\boldsymbol{w}\|^2}\boldsymbol{w} \tag{54}$$

向量 $\boldsymbol{d}$ 模便是直线/平面外一点 $\boldsymbol{q}$ 到直线 (51) 的距离 d：

$$d = \|\boldsymbol{d}\| = \frac{|\boldsymbol{w}^{\mathrm{T}}\boldsymbol{a}|}{\|\boldsymbol{w}\|} = \frac{|\boldsymbol{w}\cdot\boldsymbol{a}|}{\|\boldsymbol{w}\|} \tag{55}$$

考虑到 $\boldsymbol{w}^{\mathrm{T}}\boldsymbol{a}$ 结果为标量，因此 (55) 分子仅用绝对值符号。

将 (53) 代入 (55)，整理得到：

$$d = \frac{|\boldsymbol{w}^{\mathrm{T}}(\boldsymbol{q}-\boldsymbol{x})|}{\|\boldsymbol{w}\|} = \frac{|\boldsymbol{w}^{\mathrm{T}}\boldsymbol{q}-\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}|}{\|\boldsymbol{w}\|} = \frac{|\boldsymbol{w}\cdot\boldsymbol{q}-\boldsymbol{w}\cdot\boldsymbol{x}|}{\|\boldsymbol{w}\|} \tag{56}$$

将 (52) 代入 (56) 得到：

$$d = \frac{|\boldsymbol{w}^{\mathrm{T}}(\boldsymbol{q}-\boldsymbol{x})|}{\|\boldsymbol{w}\|} = \frac{|\boldsymbol{w}^{\mathrm{T}}\boldsymbol{q}+b|}{\|\boldsymbol{w}\|} = \frac{|\boldsymbol{w}\cdot\boldsymbol{q}+b|}{\|\boldsymbol{w}\|} \tag{57}$$

正交投影点坐标

下面求解点 $\boldsymbol{q}$ 在超平面上的正交投影点为 $\boldsymbol{x}_q$。

$\boldsymbol{x}_q$ 在超平面上，因此下式成立：

$$\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_q + b = 0 \tag{58}$$

此外，$\boldsymbol{w}$ 平行于 $\boldsymbol{x}_q - \boldsymbol{q}$ 由此可以构造第二个等式：

$$\boldsymbol{x}_q - \boldsymbol{q} = k\boldsymbol{w} \tag{59}$$

k 为任意非零实数。整理上式，$\boldsymbol{x}_q$ 为：

$$\boldsymbol{x}_q = k\boldsymbol{w} + \boldsymbol{q} \tag{60}$$

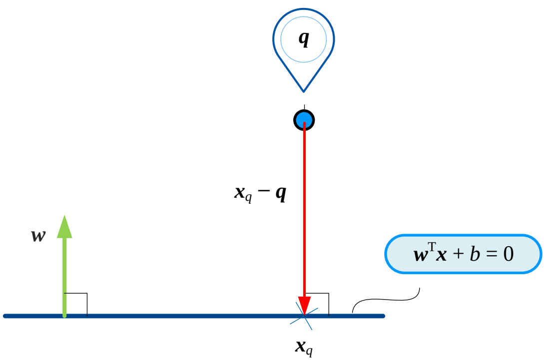


图 17. 直线外一点到直线的正交投影点

将 (60) 代入 (58)，得到：

$$\boldsymbol{w}^{\mathrm{T}}(k\boldsymbol{w}+\boldsymbol{q}) + b = 0 \tag{61}$$

整理上式得到 k

$$k = -\frac{\left(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{q}+b\right)}{\boldsymbol{w}^{\mathrm{T}}\boldsymbol{w}} \tag{62}$$

将上式代入 (60)，得到正交投影点为 $\boldsymbol{x}_q$：

$$\boldsymbol{x}_q = \boldsymbol{q} - \frac{\left(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{q}+b\right)}{\boldsymbol{w}^{\mathrm{T}}\boldsymbol{w}}\boldsymbol{w} \tag{63}$$

向量在平面内投影

同样利用上述投影思路，可以计算如图 18 所示的向量 $\boldsymbol{q}$ 在平面 H ($\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} + b = 0$) 的投影：

$$\mathrm{proj}_H\left(\boldsymbol{q}\right) = \boldsymbol{q} - \mathrm{proj}_{\boldsymbol{w}}\left(\boldsymbol{q}\right) = \boldsymbol{q} - \frac{\boldsymbol{w}^{\mathrm{T}}\boldsymbol{q}}{\boldsymbol{w}^{\mathrm{T}}\boldsymbol{w}}\boldsymbol{w} \tag{64}$$

比较 (63) 和 (64)，可以发现 (64) 就是 (63) 中 $b = 0$ 的特殊情况。

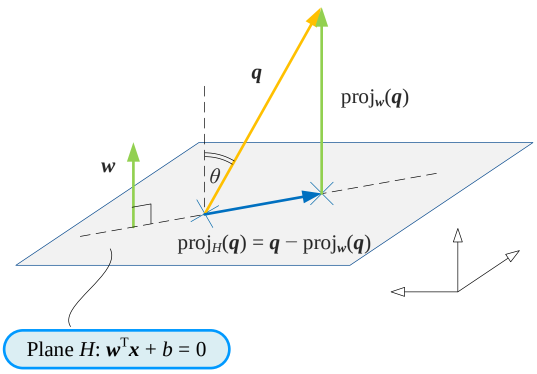


图 18. 向量 $\boldsymbol{q}$ 在平面 H 的投影

平行面距离

给定两个相互平行超平面的解析式如下：

$$\begin{cases} \boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} + b_1 = 0 \\ \boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} + b_2 = 0 \end{cases} \tag{65}$$

如图 19 所示，A 和 B 分别位于这两个超平面上，A 点坐标为 $\boldsymbol{x}_A$，B 点坐标为 $\boldsymbol{x}_B$。构造如下等式：

$$\begin{cases} \boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_A + b_1 = 0 \\ \boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_B + b_2 = 0 \end{cases} \Rightarrow \begin{cases} \boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_A = -b_1 \\ \boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_B = -b_2 \end{cases} \tag{66}$$

构造向量 $\boldsymbol{a}$ 起点为 B，终点为 A：

$$\boldsymbol{a} = \boldsymbol{x}_A - \boldsymbol{x}_B \tag{67}$$

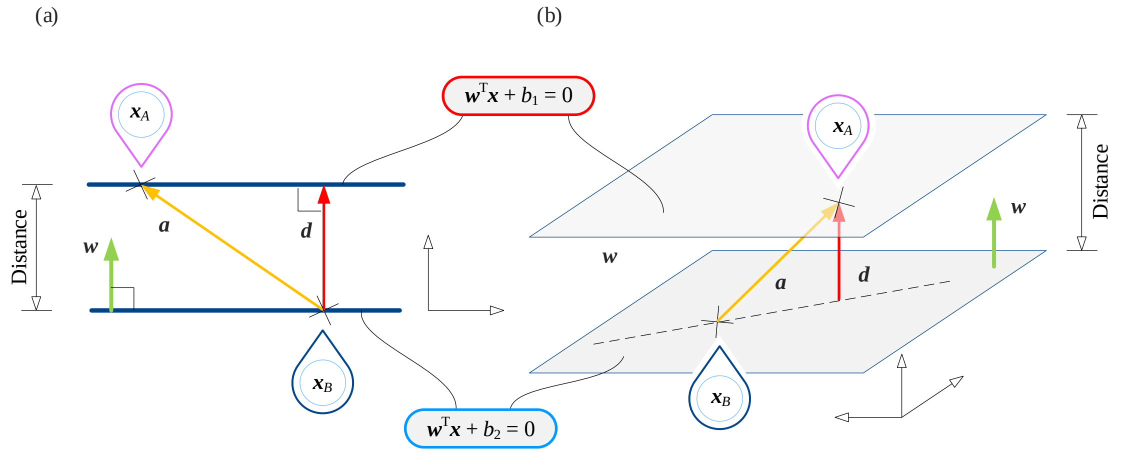


图 19. 利用向量投影计算间隔宽度

根据 (55)，平行超平面距离为向量 $\boldsymbol{a}$ 在直线梯度向量 $\boldsymbol{w}$ 上的投影：

$$\frac{\left|\boldsymbol{w}^{\mathrm{T}}\boldsymbol{a}\right|}{\|\boldsymbol{w}\|}=\frac{|\boldsymbol{w}\cdot\boldsymbol{a}|}{\|\boldsymbol{w}\|}=\frac{\left|\boldsymbol{w}^{\mathrm{T}}\left(\boldsymbol{x}_A-\boldsymbol{x}_B\right)\right|}{\|\boldsymbol{w}\|}=\frac{\left|-b_1-\left(-b_2\right)\right|}{\|\boldsymbol{w}\|}=\frac{\left|b_2-b_1\right|}{\|\boldsymbol{w}\|} \tag{68}$$

相比本书之前内容，本章内容很特殊。本章之前在讲解线性代数工具时，我们经常利用几何视角。而本章正好相反，本章讲解的是几何知识，采用的是线性代数工具。

有向量的地方，就有几何！

本章内容告诉我们，这句话反过来也正确。有几何的地方，就有向量。

本书讲解的几何知识对于很多机器学习、数据科学算法非常重要。本系列丛书在讲到具体算法时，会提醒大家其中用到了哪些本章和下两章介绍的几何知识。

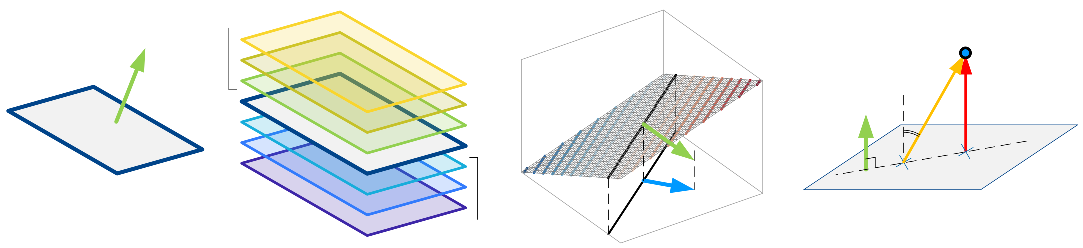


图 20. 总结本章重要内容的四副图