



最优化方法：数学基础回顾

向量与矩阵、导数与微分、泰勒展开与二次型

梁经纬 Jingwei Liang

Institute of Natural Sciences
School of Mathematical Sciences

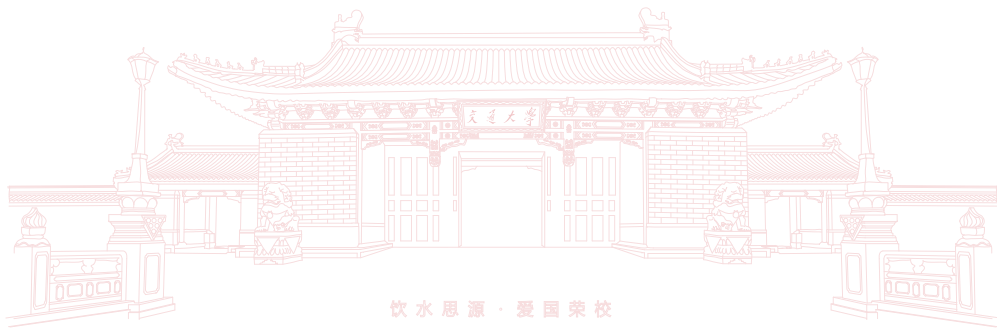
Lecture 00: 数学基础

2026 年秋季学期

饮水思源 · 爱国荣校

01

课程信息



饮水思源 · 爱国荣校

课程安排

授课与答疑

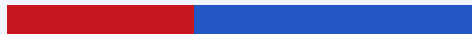
授课教师：梁经纬

周三 14:00–16:00, 理科楼 6 号楼 355 室

成绩构成

平时分 40%

期末考试 60%



主要参考文献

Boyd · Vandenberghe
Convex Optimization

Cambridge University Press, 2004



参考文献

建模与数值方法

- ▣ (刘浩洋、户将、李勇锋、文再文)
(《最优化：建模、算法与理论》), 高等教育出版社。
- ▣ S. Boyd and L. Vandenberghe
Convex Optimization, Cambridge University Press, 2004.
- ▣ A. Beck
First-order Methods in Optimization, MOS-SIAM Series on Optimization, 2017.
- ▣ J. Nocedal and S. Wright
Numerical Optimization, Springer Science & Business Media, 2006.

优化理论与凸分析

- ▣ B. Polyak
Introduction to Optimization, Optimization Software, 1987.
- ▣ Y. Nesterov
Introductory Lectures on Convex Optimization: A Basic Course, Vol. 87, Springer, 2013.
- ▣ H. H. Bauschke and P. L. Combettes
Convex Analysis and Monotone Operator Theory in Hilbert Spaces, Vol. 408, Springer, 2011.
- ▣ R. T. Rockafellar
Convex Analysis, Princeton University Press, 2015.

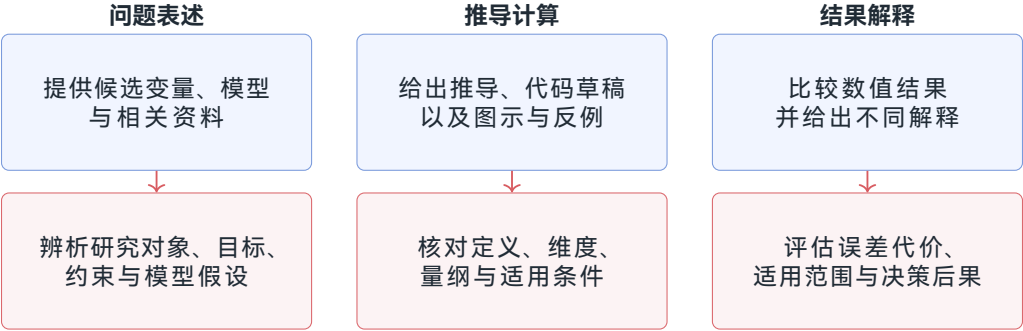
阅读线索



数学工作的分工

2025 年，特定版本的 Gemini Deep Think 在国际数学奥林匹克竞赛试题上达到金牌标准。对于明确给定的问题，大模型已经能完成相当复杂的推导。

大模型可提供候选结果 数学学习要形成判断依据

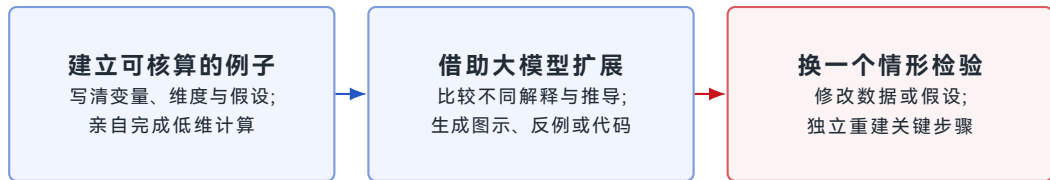


推导正确只说明表达式内部成立；模型是否对应真实任务，还取决于对问题和结果的判断。

Google DeepMind, 2025: IMO gold-medal standard

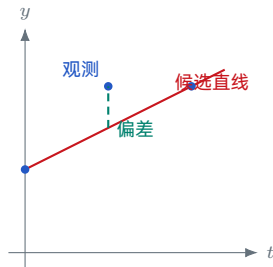
数学学习方式

基本概念和关键运算仍要亲自掌握；重复计算、图示与多种推导可以借助工具完成。两者共同服务于对问题的理解。



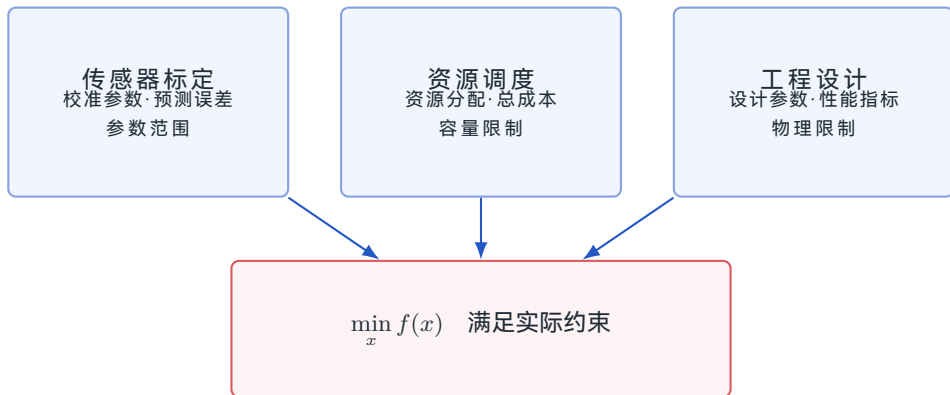
以时间 t (秒) 与位移 y (米) 的测量为例，工具可以提出直线模型并作图。我们仍需解释 $y \approx a + bt$ 中各量的单位与测量偏差。

问题 图中的偏差能否接受？还需要知道什么？



为什么学优化？

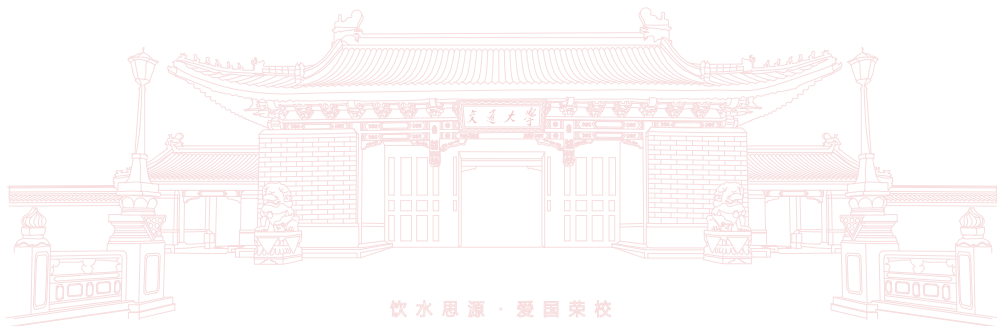
不同专业的任务各异，但都需要把**可选择的变量**、**评价标准**和**实际限制**联系起来。



本课程继续研究：**怎样建立可信的优化模型**，**如何判断解的性质**，以及**如何选择并评估求解方法**。向量和矩阵、导数与二阶近似，是这些讨论的共同语言。

02

向量与矩阵



饮水思源 · 爱国荣校

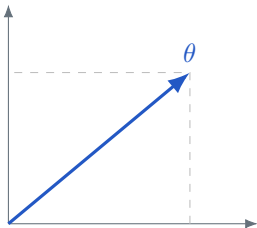
数学基础回顾

从测量数据与简单函数出发，连接三组数学工具。

向量与矩阵

变量、关系与误差

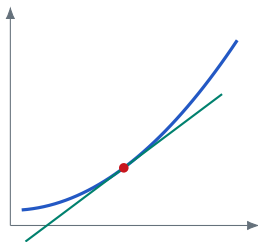
$$\theta, \quad A\theta - y$$



导数与微分

函数的一阶变化

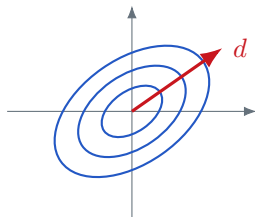
$$df(x)[d] = \nabla f(x)^\top d$$



泰勒展开与二次型

函数的局部二阶形状

$$\frac{1}{2}d^\top H d$$



变量的表示 $\rightarrow$ 函数的变化 $\rightarrow$ 局部近似

直线与测量数据

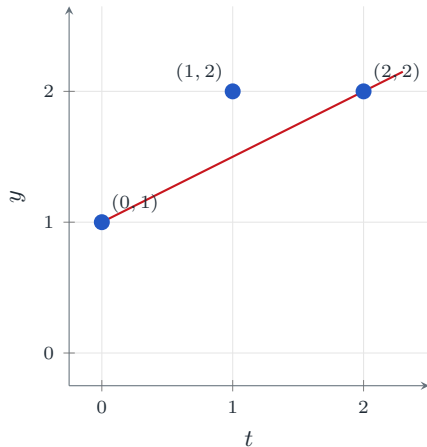
测得三组数据:

输入 t	0	1	2
观测值 y	1	2	2

设模型为 $y \approx a + bt$, 对应的三个预测值是

$$a, \quad a + b, \quad a + 2b.$$

数据 t, y 已给定; **参数** a, b 待确定。



候选模型: $y = 1 + 0.5t$

参数与向量

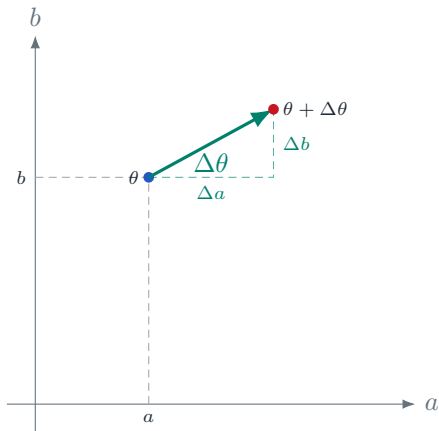
把两个参数合成列向量：

$$\theta = \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2.$$

同时改变两个参数时，

$$\theta + \Delta\theta = \begin{pmatrix} a + \Delta a \\ b + \Delta b \end{pmatrix}.$$

同一组坐标可表示**点**；作为增量时，表示**位移**。



数乘与线性组合

设

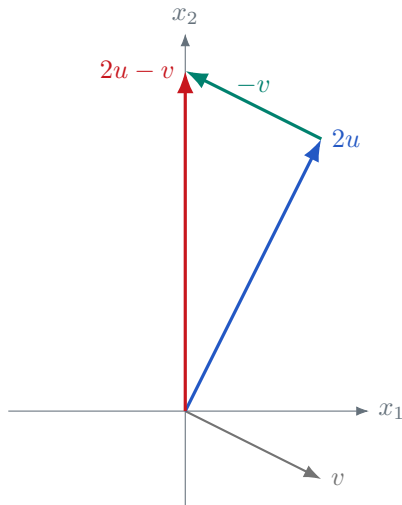
$$u = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad v = \begin{pmatrix} 2 \\ -1 \end{pmatrix}.$$

逐分量计算得到

$$2u - v = \begin{pmatrix} 2 \\ 4 \end{pmatrix} - \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 5 \end{pmatrix}.$$

标准坐标向量 $e_1 = (1, 0)^\top$ 、 $e_2 = (0, 1)^\top$ 满足

$$(a, b)^\top = ae_1 + be_2.$$



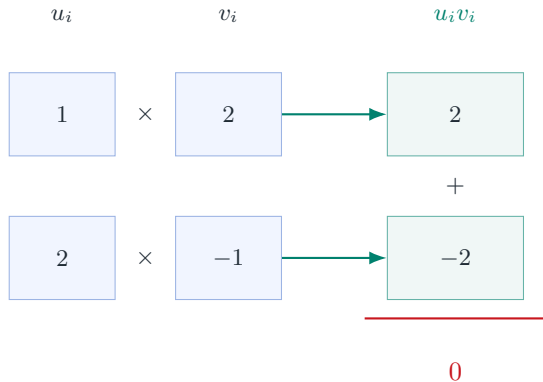
内积

欧氏内积 (inner product)

$$u^{\top} v = \sum_{i=1}^n u_i v_i, \quad u, v \in \mathbb{R}^n.$$

例如

$$u = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad v = \begin{pmatrix} 2 \\ -1 \end{pmatrix}.$$



对应分量相乘，再求和。

$$\underbrace{u^{\top}}_{1 \times n} \underbrace{v}_{n \times 1} = \underbrace{u^{\top} v}_{1 \times 1}$$

输入分量数相同，输出为**标量**。

夹角与投影

非零向量夹角为 ϑ 时,

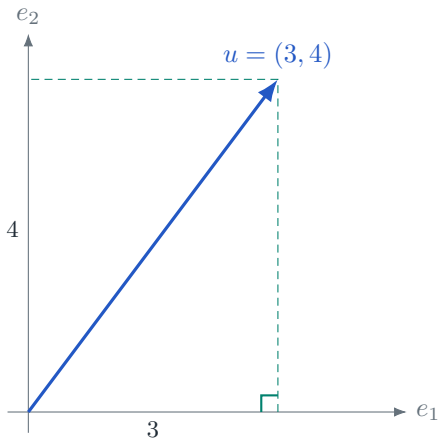
$$u^\top v = \|u\|_2 \|v\|_2 \cos \vartheta.$$

因此内积的符号对应锐角、直角与钝角;

$u^\top v = 0$ 表示正交 (orthogonal)。

若 e 是单位向量, $u^\top e$ 是 u 沿 e 的**有符号投影长度**。

$$u = (3, 4)^\top : \quad u^\top e_1 = 3, \quad u^\top e_2 = 4.$$



范数与距离

欧氏范数 (Euclidean norm) 为

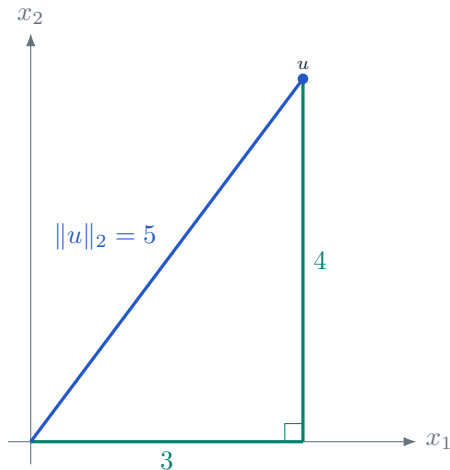
$$\|u\|_2 = \sqrt{u^\top u} = \sqrt{u_1^2 + \cdots + u_n^2}.$$

两个点 p, q 的距离是 $\|p - q\|_2$ 。

对于 $u = (3, 4)^\top$,

$$\|u\|_2 = 5, \quad \|u\|_2^2 = 25.$$

问题 若 $w = (6, 8)^\top$, $\|w\|_2$ 与 $\|w\|_2^2$ 各是多少?

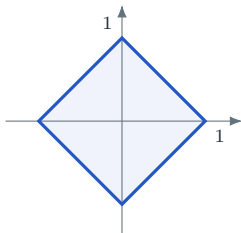


三种向量范数

ℓ_1 范数

绝对值之和

$$\|u\|_1 = \sum_i |u_i|$$

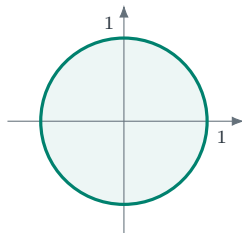


$$\|(3, 4)^T\|_1 = 7$$

ℓ_2 范数

欧氏长度

$$\|u\|_2 = \sqrt{\sum_i u_i^2}$$

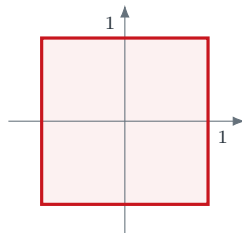


$$\|(3, 4)^T\|_2 = 5$$

ℓ_∞ 范数

绝对值的最大值

$$\|u\|_\infty = \max_i |u_i|$$



$$\|(3, 4)^T\|_\infty = 4$$

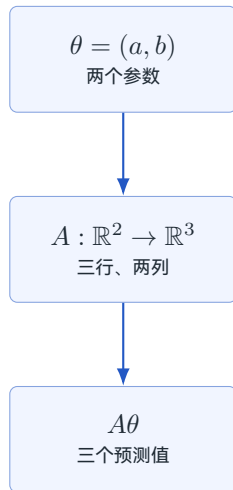
图中边界对应 $\|u\| = 1$ ，内部对应 $\|u\| \leq 1$ 。后续主要使用欧氏范数。

矩阵与预测

$$\underbrace{\begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}}_{A \in \mathbb{R}^{3 \times 2}} \underbrace{\begin{pmatrix} a \\ b \end{pmatrix}}_{\theta \in \mathbb{R}^2} = \underbrace{\begin{pmatrix} a \\ a+b \\ a+2b \end{pmatrix}}_{A\theta \in \mathbb{R}^3}.$$

列数 = 输入参数数； 行数 = 输出预测值数。

输入 t	0	1	2
对应矩阵行	$[1 \ 0]$	$[1 \ 1]$	$[1 \ 2]$



矩阵乘法：按行

A 的每一行与参数向量取内积，得到一个预测值：

$$(A\theta)_i = \sum_{j=1}^2 A_{ij}\theta_j.$$

输入

行向量 $\times$ 参数

预测

$t = 0$

$$\begin{bmatrix} 1 & 0 \end{bmatrix} \begin{pmatrix} a \\ b \end{pmatrix}$$

a

$t = 1$

$$\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{pmatrix} a \\ b \end{pmatrix}$$

$a + b$

$t = 2$

$$\begin{bmatrix} 1 & 2 \end{bmatrix} \begin{pmatrix} a \\ b \end{pmatrix}$$

$a + 2b$

矩阵乘法：按列

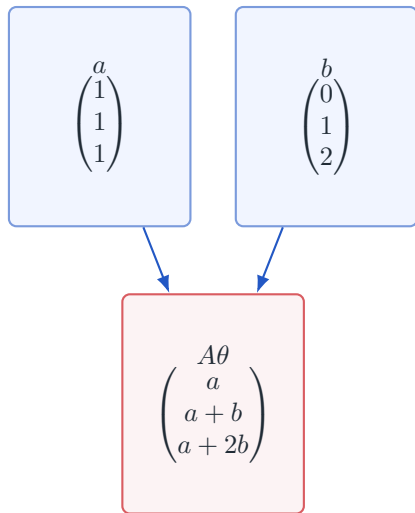
设矩阵列为 $a_1, \dots, a_n$ ，则

$$Ax = x_1 a_1 + \dots + x_n a_n.$$

本例：

$$A\theta = a \underbrace{\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}}_{\text{第一列}} + b \underbrace{\begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}}_{\text{第二列}}.$$

输出由列向量的线性组合构成。



矩阵同时定义线性映射 $A : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ 。

转置与维度

转置 (transpose) 交换行与列:

$$(A^T)_{ji} = A_{ij}.$$

矩阵相乘时, **内侧维度相等**:

$$(m \times n)(n \times p) \longrightarrow m \times p.$$

问题 $A^T A$ 和 AA^T 各是几乘几?

1	0
1	1
1	2

$\xrightarrow{T}$

1	1	1
0	1	2

$$A : 3 \times 2$$

$$A^T : 2 \times 3$$

$$A^T A : (2 \times 3)(3 \times 2) \rightarrow 2 \times 2$$

$$AA^T : (3 \times 2)(2 \times 3) \rightarrow 3 \times 3$$

乘法顺序

对 $u = (1, 2)^\top$ 、 $v = (2, -1)^\top$:

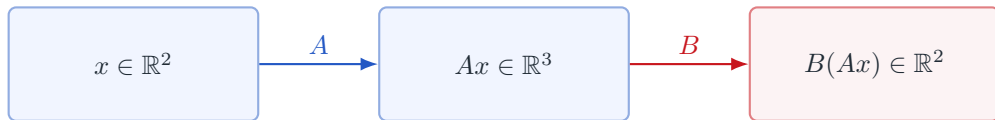
内积: 标量

$$\underbrace{u^\top v}_{(1 \times 2)(2 \times 1)} = 0.$$

外积: 矩阵

$$\underbrace{uv^\top}_{(2 \times 1)(1 \times 2)} = \begin{pmatrix} 2 & -1 \\ 4 & -2 \end{pmatrix}.$$

复合映射中, 最右侧的矩阵先作用:



$$(AB)^\top = B^\top A^\top, \quad AB \text{ 通常不等于 } BA.$$

线性无关与秩

列向量 $a_1, \dots, a_n$ 线性无关, 指

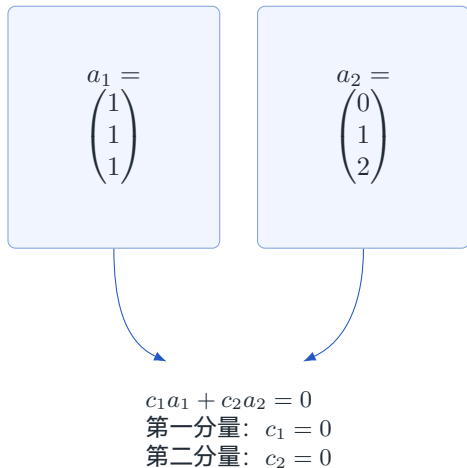
$$c_1 a_1 + \dots + c_n a_n = 0 \implies c_1 = \dots = c_n = 0.$$

秩 (rank) 是线性无关列的最大个数。

本例 A 的两列线性无关, 所以

$$\text{rank}(A) = 2.$$

列数也为 2, 称为**满列秩**。



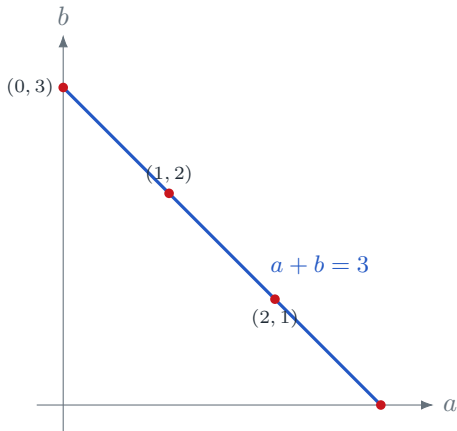
秩与参数识别

若三次输入均为 $t = 1$,

$$\tilde{A} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad \tilde{A} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a+b \\ a+b \\ a+b \end{pmatrix}.$$

两列相同, $\text{rank}(\tilde{A}) = 1$ 。

$(1, 2)$ 、 $(2, 1)$ 、 $(0, 3)$ 产生相同输出 $(3, 3, 3)^\top$ 。



零空间

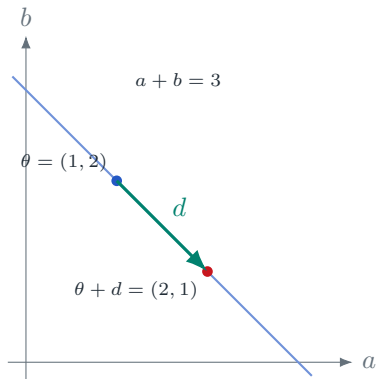
零空间 (null space) 包含不改变输出的参数增量:

$$\ker A = \{d \in \mathbb{R}^n : Ad = 0\}.$$

对秩不足的 $\tilde{A}$,

$$d = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \tilde{A}d = 0.$$

$$\tilde{A}(\theta + sd) = \tilde{A}\theta, \quad s \in \mathbb{R}.$$



$$\theta = (1, 2)^\top \quad \text{与} \quad \theta + d = (2, 1)^\top \quad \xrightarrow{\tilde{A}} \quad (3, 3, 3)^\top$$

方程的可解性

按列理解 $Ax = c$:

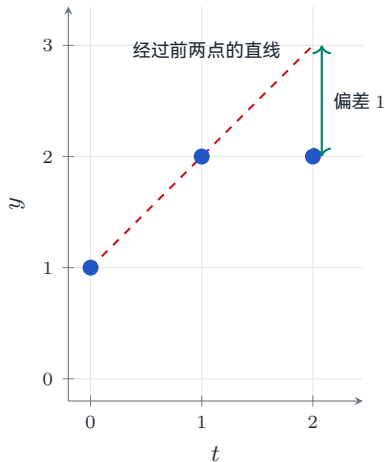
$$c = x_1 a_1 + \cdots + x_n a_n.$$

有解，当且仅当 $c \in \text{col}(A)$; **列空间** $\text{col}(A)$ 由各列的线性组合构成。

对原直线模型，任意预测值 $z = A\theta$ 满足

$$z_1 - 2z_2 + z_3 = 0.$$

观测 $y = (1, 2, 2)^\top$ 却给出 $1 - 2(2) + 2 = -1$ ，所以 $A\theta = y$ 无解。



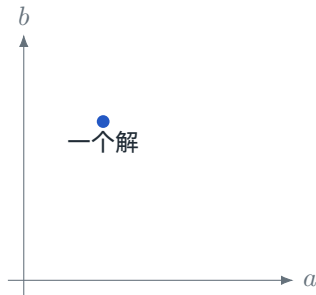
解的唯一性

若 x_p 是 $Ax = c$ 的一个解, 则全部解为

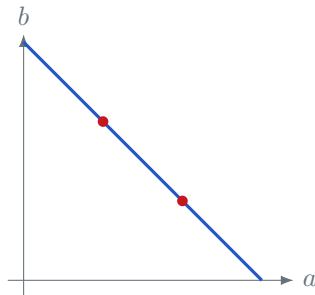
$$x_p + \ker A.$$

因此在**已有解**的前提下,

$$\text{解唯一} \iff \ker A = \{0\}.$$



$$c = A(1, 2)^T, \ker A = \{0\}$$



$$\tilde{A}: c = (3, 3, 3)^T$$

残差与平方和

预测与观测的差称为残差 (residual):

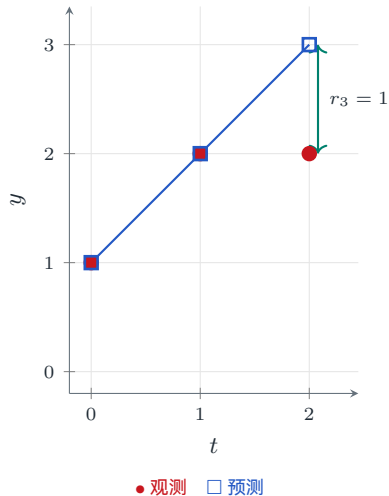
$$r = A\theta - y, \quad \|r\|_2^2 = r^\top r = \sum_{i=1}^3 r_i^2.$$

例如 $\theta = (1, 1)^\top$ 时,

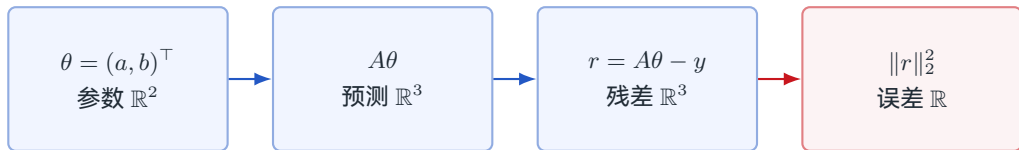
$$A\theta = (1, 2, 3)^\top, \quad y = (1, 2, 2)^\top, \quad r = (0, 0, 1)^\top.$$

所以 $\|r\|_2^2 = 1$ 。任务可写作

$$\min_{\theta \in \mathbb{R}^2} \|A\theta - y\|_2^2.$$



从参数到误差



表达式

含义

$$x \in \mathbb{R}^n$$

n 个变量组成一个列向量。

$$Ax$$

多个线性关系；列向量的线性组合。

$$u^\top v, \|u\|_2$$

两个向量的内积；一个向量的欧氏长度。

$$Ad = 0$$

参数变化 d 不改变输出。

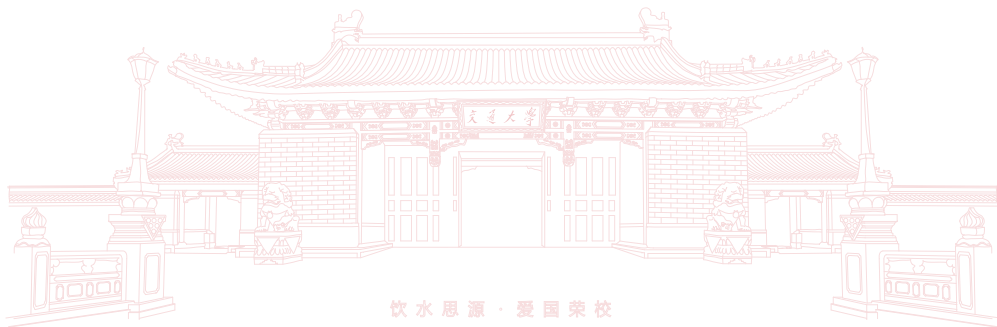
$$\|Ax - c\|_2^2$$

预测与给定数据之间的误差平方和。

输入向量改变后，函数值怎样变化？ $\longrightarrow$ 导数与微分

03

导数与微分



饮水思源 · 爱国荣校

多元函数

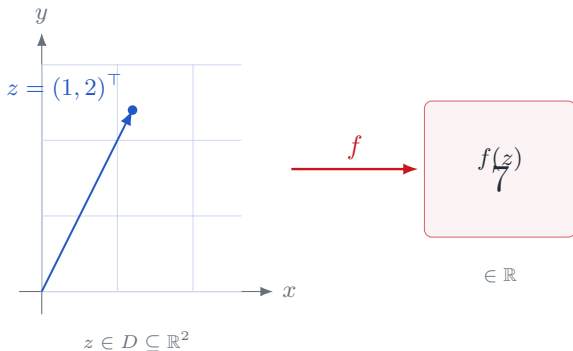
参数组成向量；成本、误差或能耗可以表示为实值函数：

$$f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}.$$

本节使用

$$f(x, y) = x^2 + xy + y^2, \quad z = (x, y)^\top.$$

例如 $f(1, 2) = 7$ 。向量 z 是输入，函数值 $f(z)$ 是输出；定义域 D 给出允许代入的输入。



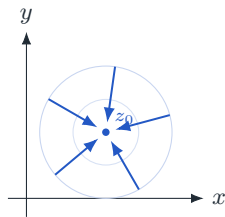
连续性

一元函数在 t_0 连续:

$$\lim_{h \rightarrow 0} g(t_0 + h) = g(t_0).$$

多元函数在 z_0 连续:

$$\lim_{\|d\|_2 \rightarrow 0} f(z_0 + d) = f(z_0).$$



任意允许方向都接近 z_0

连续表示输出没有突变。多元极限必须对**所有**允许的接近方式成立，而不只是沿坐标轴接近。

连续性并不自动保证导数存在: $g(t) = |t|$ 在 0 连续, 但左、右导数分别为 -1 与 1 。

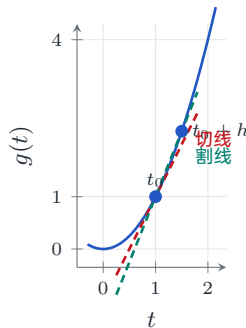
一元导数

$$g'(t_0) = \lim_{h \rightarrow 0} \frac{g(t_0 + h) - g(t_0)}{h}.$$

对 $g(t) = t^2$,

$$\frac{(t_0 + h)^2 - t_0^2}{h} = 2t_0 + h \rightarrow 2t_0.$$

可导 $\implies$ **连续**; 反向一般不成立。



一阶近似

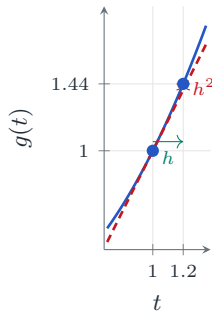
当 h 很小时,

$$g(t_0 + h) \approx g(t_0) + g'(t_0)h.$$

在 $t_0 = 1$,

$$(1+h)^2 = \underbrace{1}_{\text{原值}} + \underbrace{2h}_{\text{一阶变化}} + \underbrace{h^2}_{\text{余项}}.$$

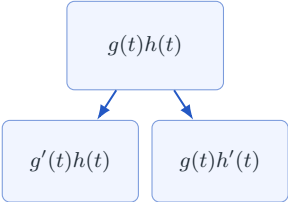
取 $h = 0.01$: 线性近似为 1.02, 精确值为 1.0201。



图示取 $h = 0.2$, 以便看清二阶偏差。

基本求导规则

函数或运算	导数
常数 c	0
t^p (正整数 p)	pt^{p-1}
e^t	e^t
$\log t$ ($t > 0$)	$1/t$
$ag(t) + bh(t)$	$ag'(t) + bh'(t)$
$g(t)h(t)$	$g'(t)h(t) + g(t)h'(t)$



两个因子的局部变化相加

乘积求导一般不是 $g'(t)h'(t)$ 。

偏导数

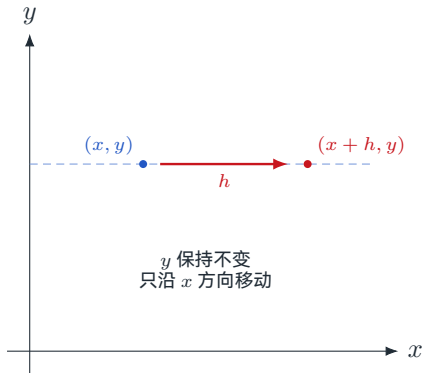
对二元函数 $f(x, y)$,

$$f_x(x, y) = \lim_{h \rightarrow 0} \frac{f(x + h, y) - f(x, y)}{h}.$$

求 f_x 时, 把 y 暂时看作常数。例如

$$\frac{\partial}{\partial x}(xy) = y, \quad \frac{\partial}{\partial x}(y^2) = 0.$$

保持 y 不变, 并不意味着把 y 设为零。



偏导数计算

$$f(x, y) = x^2 + xy + y^2.$$

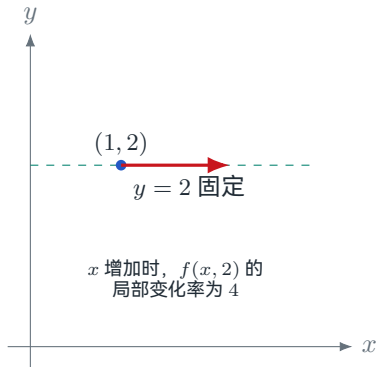
保持 y 不变:

$$f_x(x, y) = 2x + y.$$

因此

$$f_x(1, 2) = 4.$$

问题 求 $f_y(x, y)$ 与 $f_y(1, 2)$ 。



梯度

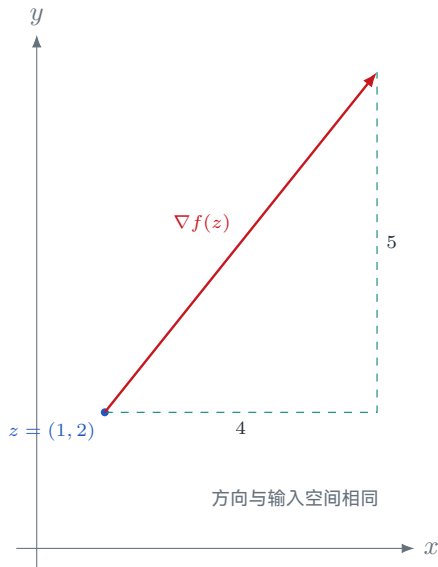
对 $f: \mathbb{R}^n \rightarrow \mathbb{R}$,

$$\nabla f(z) = \begin{pmatrix} \frac{\partial f}{\partial z_1}(z) \\ \vdots \\ \frac{\partial f}{\partial z_n}(z) \end{pmatrix} \in \mathbb{R}^n.$$

本例中

$$\nabla f(x, y) = \begin{pmatrix} 2x + y \\ x + 2y \end{pmatrix}, \quad \nabla f(1, 2) = \begin{pmatrix} 4 \\ 5 \end{pmatrix}.$$

梯度的维数与输入的维数相同。



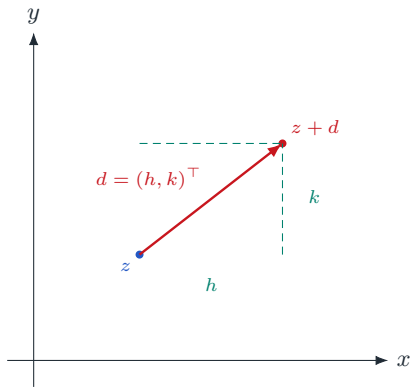
全微分

令 $z = (x, y)^\top$ 、 $d = (h, k)^\top$ 。当 x, y 分别改变 h, k ，可微函数满足

$$f(x + h, y + k) - f(x, y) \approx f_x(x, y)h + f_y(x, y)k.$$

$$f(z + d) - f(z) \approx \nabla f(z)^\top d.$$

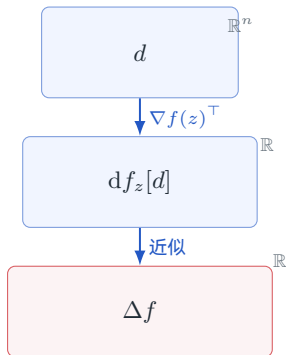
即为**梯度与增量的内积**。



增量、梯度与微分

固定基点 z 后:

对象	表达式	类型
输入增量	d	向量
梯度	$\nabla f(z)$	向量
实际函数增量	$\Delta f = f(z + d) - f(z)$	标量
微分作用于增量	$df_z[d] = \nabla f(z)^\top d$	标量



传统写法: $df = f_x dx + f_y dy$ 。微分是关于输入增量的线性函数, 通常不等于实际增量。

可微性

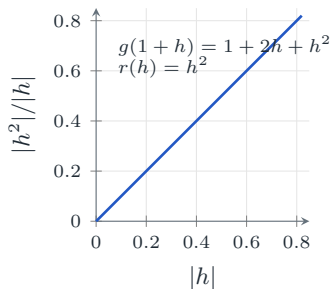
在内点 z ，可微意味着存在向量 g ，使得

$$f(z+d) = f(z) + g^\top d + r(d),$$

$$\frac{|r(d)|}{\|d\|_2} \longrightarrow 0 \quad (d \rightarrow 0).$$

此时 $g = \nabla f(z)$ 。同一个线性近似适用于所有足够小的增量 d 。

偏导数在邻域内存在且在该点连续，是可微的常用充分条件。



一元例子中，相对误差 $|h|$ 随 $h \rightarrow 0$ 而消失。

微分估计

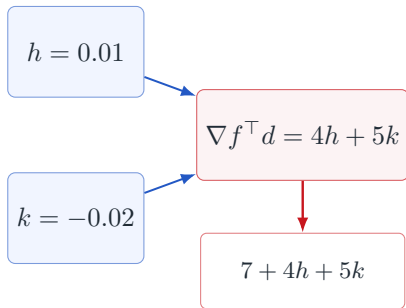
仍取 $f(x, y) = x^2 + xy + y^2$ 。在 $z = (1, 2)^\top$:

$$f(z) = 7, \quad \nabla f(z) = \begin{pmatrix} 4 \\ 5 \end{pmatrix}.$$

对新点 $(1.01, 1.98)$,

$$d = \begin{pmatrix} 0.01 \\ -0.02 \end{pmatrix}.$$

问题 用 $f(z + d) \approx f(z) + \nabla f(z)^\top d$ 估计 $f(1.01, 1.98)$ 。



切平面

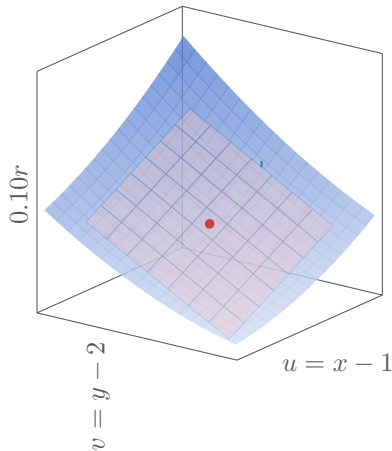
在 (x_0, y_0) 附近,

$$\begin{aligned} f(x, y) &\approx f(x_0, y_0) \\ &\quad + f_x(x_0, y_0)(x - x_0) \\ &\quad + f_y(x_0, y_0)(y - y_0). \end{aligned}$$

对 $f = x^2 + xy + y^2$, 在 $(1, 2)$ 处的切平面是

$$w = 7 + 4(x - 1) + 5(y - 2).$$

一元函数使用切线; 二元函数使用切平面。



$r = f(x, y) - 7$; 竖轴按 0.10 缩放

蓝色: 曲面 红色: 局部切平面

方向导数

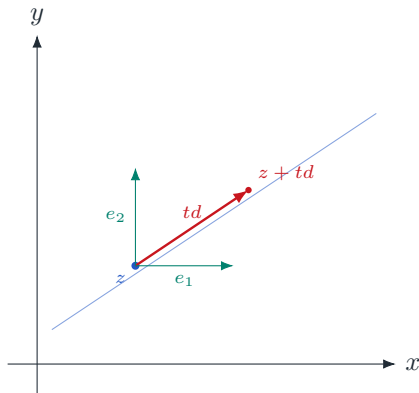
给定单位向量 d , 沿直线 $z + td$ 观察:

$$\varphi(t) = f(z + td).$$

若 f 在 z 处可微,

$$D_d f(z) = \lim_{t \rightarrow 0} \frac{f(z + td) - f(z)}{t} = \nabla f(z)^\top d.$$

偏导数对应方向 $e_1, \dots, e_n$; 方向导数允许任意指定的单位方向。



方向导数的计算

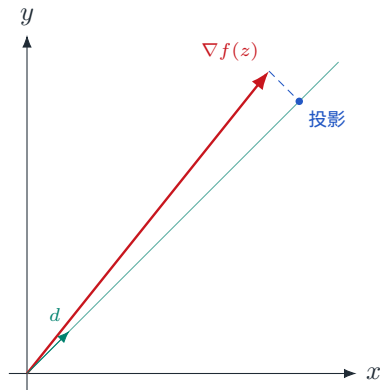
在 $z = (1, 2)^\top$,

$$\nabla f(z) = \begin{pmatrix} 4 \\ 5 \end{pmatrix}.$$

沿坐标方向的变化率分别为 4 与 5。沿单位方向

$$d = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

的变化率是多少？



链式法则

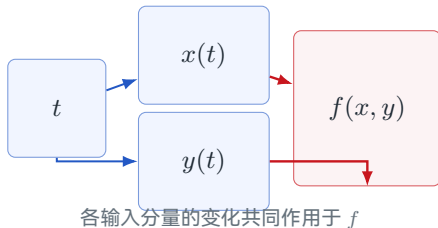
若 $z(t) = (x(t), y(t))^{\top}$ 可导, 且 f 可微,

$$\begin{aligned}\frac{d}{dt}f(x(t), y(t)) &= f_x(x(t), y(t))x'(t) \\ &\quad + f_y(x(t), y(t))y'(t).\end{aligned}$$

向量形式:

$$\frac{d}{dt}f(z(t)) = \nabla f(z(t))^{\top} z'(t).$$

函数的变化率由各输入分量的变化速度共同决定。



沿路径求导

设

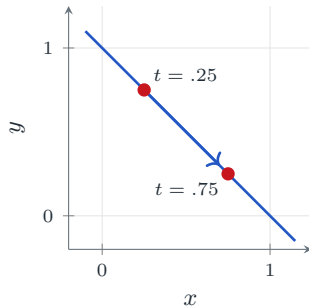
$$f(x, y) = x^2 + xy + y^2, \quad x(t) = t, \quad y(t) = 1 - t.$$

先代入:

$$\begin{aligned} f(t, 1 - t) &= t^2 + t(1 - t) + (1 - t)^2 \\ &= t^2 - t + 1. \end{aligned}$$

因此

$$\frac{d}{dt} f(t, 1 - t) = 2t - 1.$$



链式法则计算

在 $(x, y) = (t, 1 - t)$,

$$f_x = t + 1, \quad f_y = 2 - t, \quad x'(t) = 1, \quad y'(t) = -1.$$

$$\begin{aligned} \frac{d}{dt} f(t, 1 - t) &= (t + 1) \cdot 1 + (2 - t) \cdot (-1) \\ &= 2t - 1. \end{aligned}$$

两种计算给出同一结果。

$$f_x x' = (t + 1) \cdot 1$$

$$f_y y' = (2 - t)(-1)$$

$$2t - 1$$

概念关系

连续
输入接近时，输出接近

偏导数
沿坐标方向的变化率

梯度
偏导数组成的向量

可微与微分
统一的一阶线性近似

方向导数
沿指定方向的变化率

链式法则
复合变化的计算规则

偏导数存在与可微不是同一个条件；梯度与函数增量不是同一个对象。

一阶线性近似

$$f(z + d) \approx f(z) + \nabla f(z)^\top d.$$

- ▣ 左边：新位置 $z + d$ 的函数值。
- ▣ 右边：原值加上一阶变化。
- ▣ 近似的误差相对于 $\|d\|_2$ 更小。

本例中

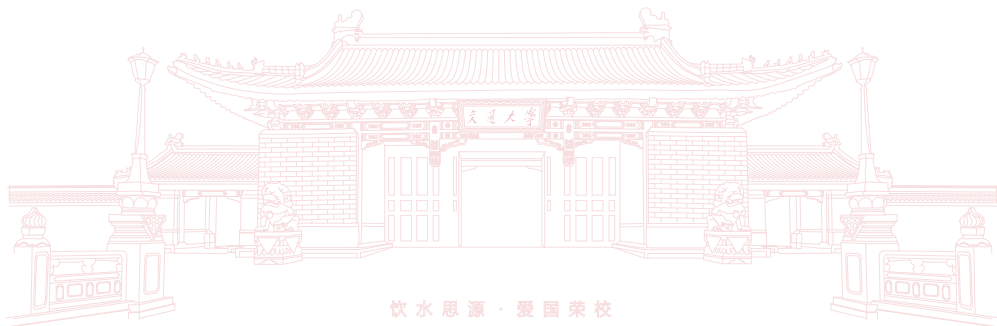
$$\begin{aligned} f(1.01, 1.98) &\approx 6.94, \\ f(1.01, 1.98) &= 6.9403. \end{aligned}$$



局部放大：显示未包含的高阶变化

04

泰勒展开与二次型



一阶近似留下的二阶误差

上一节对

$$f(x, y) = x^2 + xy + y^2$$

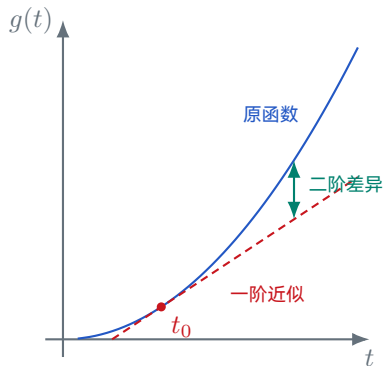
在 $z_0 = (1, 2)^\top$ 处得到一阶近似

$$f(1 + h, 2 + k) \approx 7 + 4h + 5k.$$

当 $h = 0.01, k = -0.02$ 时，线性近似为 6.94；精确值为 6.9403。

$$\underbrace{6.9403 - 6.94}_{\text{近似误差}} = 0.0003.$$

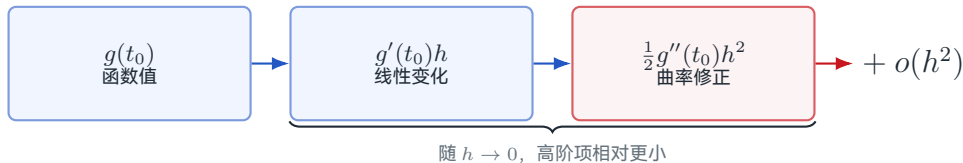
这个差异来自没有保留的二阶变化，而不是数值舍入。



一元二阶 Taylor 展开

若 g 在 t_0 附近二阶连续可导, 则

$$g(t_0 + h) = g(t_0) + g'(t_0)h + \frac{1}{2}g''(t_0)h^2 + o(h^2).$$



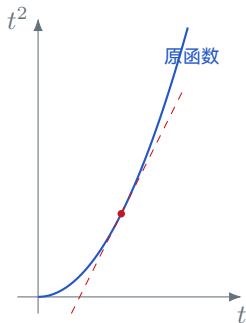
$$r(h) = o(h^2) \quad \Longleftrightarrow \quad \frac{r(h)}{h^2} \rightarrow 0 \quad (h \rightarrow 0).$$

精确展开与局部近似

$$g(t) = t^2, \quad t_0 = 1$$

$$(1+h)^2 = 1 + 2h + h^2.$$

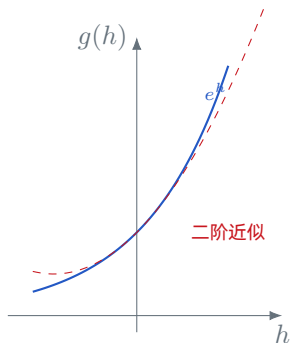
没有三阶及以上项，故上式对每个 h 都成立。



$$g(t) = e^t, \quad t_0 = 0$$

$$e^h \approx 1 + h + \frac{1}{2}h^2 \quad (h \text{ 较小时}).$$

近似多项式只描述基点附近的行为。



二元函数的二阶项

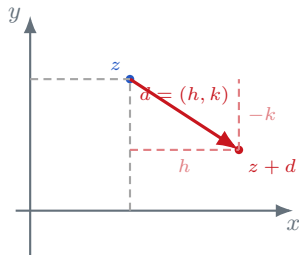
对 $f(x, y) = x^2 + xy + y^2$, 直接展开为

$$f(x+h, y+k) = f(x, y) + (2x+y)h + (x+2y)k + \underbrace{h^2 + hk + k^2}_{\text{线性近似没有保留的二阶项}}.$$

取 $h = 0.01$, $k = -0.02$, 二阶项为

$$\underbrace{h^2}_{0.0001} + \underbrace{hk}_{-0.0002} + \underbrace{k^2}_{0.0004} = 0.0003.$$

它正好等于上一页中一阶近似与精确值之差。

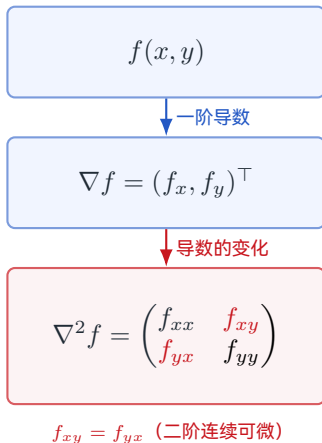


Hessian 矩阵

二阶偏导数是对一阶偏导数再求导。对二元函数，

$$\nabla^2 f(x, y) = \begin{pmatrix} f_{xx}(x, y) & f_{xy}(x, y) \\ f_{yx}(x, y) & f_{yy}(x, y) \end{pmatrix}.$$

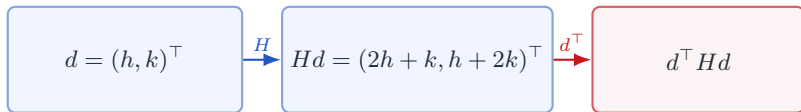
若 f 二阶连续可微，则 $f_{xy} = f_{yx}$ ，因此 Hessian 矩阵是对称矩阵。



Hessian 给出二阶项

由 $f_x = 2x + y$ 、 $f_y = x + 2y$ ，有

$$H := \nabla^2 f(x, y) = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \quad d = (h, k)^\top.$$



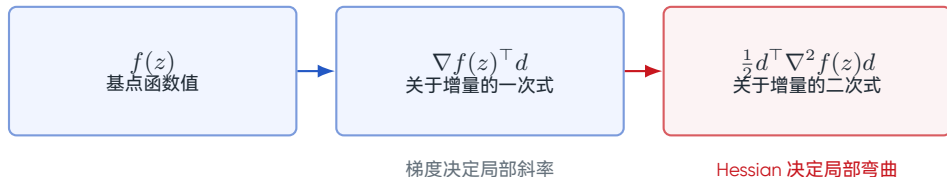
矩阵先作用于方向，再与方向作内积

$$\frac{1}{2} d^\top H d = \frac{1}{2} \begin{pmatrix} h & k \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix} = h^2 + hk + k^2.$$

多元二阶 Taylor 展开

若 f 在内点 z 附近二阶连续可微, 则

$$f(z+d) = f(z) + \nabla f(z)^\top d + \frac{1}{2}d^\top \nabla^2 f(z)d + o(\|d\|_2^2).$$



其中 $d^\top \nabla^2 f(z)d$ 的维数是 $(1 \times n)(n \times n)(n \times 1)$, 结果为标量。本例 f 是二次函数, 余项恒为零。

二次型与交叉项

对实对称矩阵

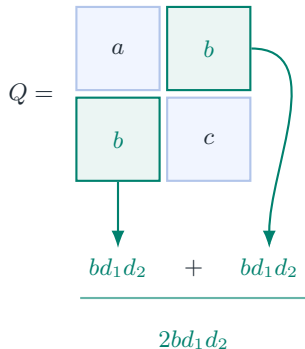
$$Q = \begin{pmatrix} a & b \\ b & c \end{pmatrix},$$

二次型为

$$q(d) = d^\top Q d = ad_1^2 + 2bd_1d_2 + cd_2^2.$$

非对角元素 b 同时在两个对称位置出现，因此交叉项系数是 $2b$ 。

问题 $d_1^2 + 6d_1d_2 + 4d_2^2$ 对应的对称矩阵是什么？



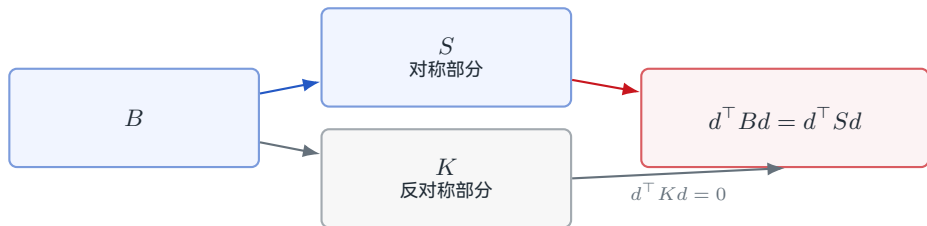
二次型只看对称部分

对任意实方阵 B ，因为 $d^\top B d$ 是标量，

$$d^\top B d = (d^\top B d)^\top = d^\top B^\top d.$$

于是可将 B 分解为对称部分与反对称部分：

$$B = \underbrace{\frac{B + B^\top}{2}}_{S=S^\top} + \underbrace{\frac{B - B^\top}{2}}_{K=-K^\top}, \quad d^\top K d = 0.$$



二次型只由矩阵的对称部分决定；以下仅讨论实对称矩阵。

二次型的符号

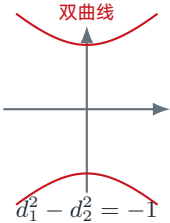
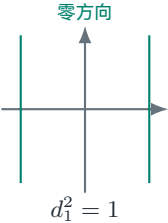
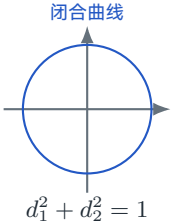
设 $Q = Q^T \in \mathbb{R}^{n \times n}$ 。

性质	对二次型 $d^T Q d$ 的要求	记号
正定	对每个 $d \neq 0$, 都有 $d^T Q d > 0$	$Q \succ 0$
半正定	对每个 d , 都有 $d^T Q d \geq 0$	$Q \succeq 0$
不定	存在方向使其为正, 也存在方向使其为负	



三个二维二次型

Q	$d^\top Q d$	性质
$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$d_1^2 + d_2^2$	正定
$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$	d_1^2	半正定但不正定
$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$	$d_1^2 - d_2^2$	不定



正定性不看元素符号

$$B = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

其所有元素为正。取 $d = (1, -1)^\top$:

$$d^\top B d = -2 < 0.$$

所以 B 不正定。

$$C = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}.$$

虽有负的非对角元素，但

$$d^\top C d = (d_1 - d_2)^2 + d_1^2 + d_2^2 > 0 \quad (d \neq 0).$$

所以 $C \succ 0$ 。

矩阵元素的正负
不能判定正定性



二次型在所有方向的符号
才是判定对象

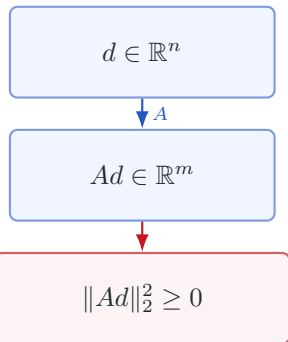
$A^\top A$ 是半正定矩阵

设 $A \in \mathbb{R}^{m \times n}$ 。对任意 $d \in \mathbb{R}^n$,

$$\begin{aligned} d^\top A^\top A d &= (Ad)^\top (Ad) \\ &= \|Ad\|_2^2 \\ &\geq 0. \end{aligned}$$

且 $A^\top A$ 是对称矩阵, 故

$$A^\top A \succeq 0.$$

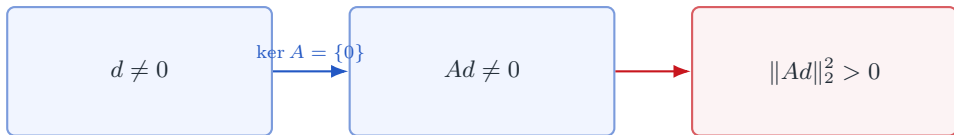


平方范数不可能为负

$A^\top A$ 何时正定?

由 $d^\top A^\top A d = \|Ad\|_2^2$ 可知, 对每个非零 d 都严格为正, 当且仅当没有非零方向被 A 映为零。因此

$$A^\top A \succ 0 \iff \ker A = \{0\} \iff A \text{ 满列秩.}$$



第一节中“参数变化是否能被输出区分”的问题, 正是这一条件的几何含义。

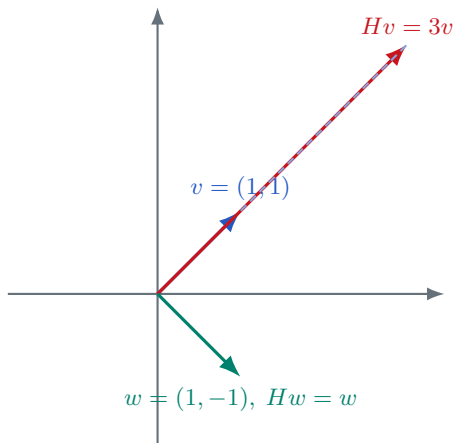
特征方向与伸缩倍数

若非零向量 v 满足 $Hv = \lambda v$, 则 v 为特征向量, λ 为特征值。

$$H = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

$$H \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 3 \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad H \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

本例的特征值为 3 和 1。



正交特征坐标

将两个特征向量单位化:

$$u_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad u_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

它们满足

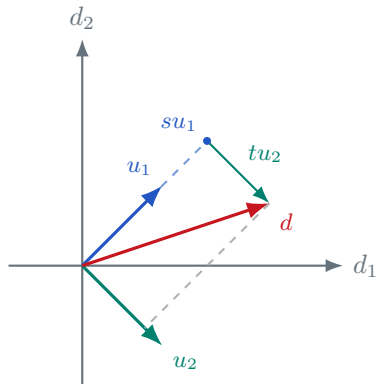
$$\|u_1\|_2 = \|u_2\|_2 = 1, \quad u_1^\top u_2 = 0.$$

任意 $d = (d_1, d_2)^\top$ 可写成

$$d = su_1 + tu_2,$$

其中新的坐标为

$$s = \frac{d_1 + d_2}{\sqrt{2}}, \quad t = \frac{d_1 - d_2}{\sqrt{2}}.$$



同一个向量 d
分别用标准坐标与特征坐标表示

特征坐标消去交叉项

由于 $Hu_1 = 3u_1$ 、 $Hu_2 = u_2$ ，对 $d = su_1 + tu_2$ ，

$$d^\top Hd = (su_1 + tu_2)^\top (3su_1 + tu_2) = 3s^2 + t^2.$$



在这组正交特征向量坐标下，交叉项消失；二次型成为平方项的加权和。

用特征值判断二次型

实对称矩阵有一组正交单位特征向量；在对应坐标中，

$$d^T Q d = \lambda_1 c_1^2 + \cdots + \lambda_n c_n^2.$$

二次型性质	特征值条件
正定	全部 $\lambda_i > 0$
半正定	全部 $\lambda_i \geq 0$
不定	同时存在正、负特征值



问题 实对称矩阵的特征值为 2, 0。它是否正定？是否半正定？

一阶项与二阶项

在基点 $z_0 = (1, 2)^\top$, 对 $d = (h, k)^\top$,

$$\begin{aligned} f(z_0 + d) &= 7 + \underbrace{(4 \quad 5) d}_{\text{一阶项}} + \underbrace{\frac{1}{2} d^\top \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} d}_{\text{二阶项}} \\ &= 7 + 4h + 5k + h^2 + hk + k^2. \end{aligned}$$

偏导数 $\rightarrow$ 梯度
内积表达一阶变化



二阶偏导数 $\rightarrow$ Hessian
二次型表达二阶变化

表达、变化与局部形状

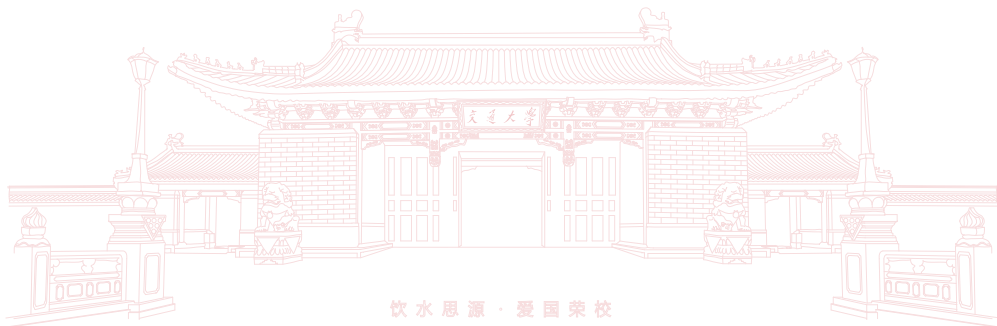
任务	核心表达
表达多个线性关系	Ax
表达大小、距离与误差	$\ x\ _2, \ Ax - c\ _2$
描述一阶变化	$\nabla f(z)^\top d$
描述二阶变化	$\frac{1}{2}d^\top \nabla^2 f(z)d$
判断二次型的符号	正定性与实对称矩阵的特征值



后续课程将在这些数学语言的基础上讨论优化模型、理论与算法。

01

补充与复习



饮水思源 · 爱国荣校

方阵、可逆与唯一解

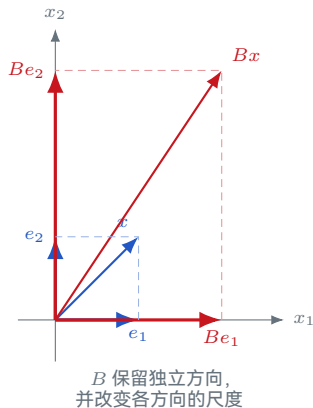
对方阵 $B \in \mathbb{R}^{n \times n}$, 下列陈述等价:

$$\begin{aligned} B \text{ 可逆} &\iff \text{rank}(B) = n \\ &\iff \ker B = \{0\} \\ &\iff \text{对每个 } b, Bx = b \text{ 有唯一解.} \end{aligned}$$

例如

$$B = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}, \quad B^{-1} = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/3 \end{pmatrix}.$$

可逆时 $B^{-1}B = BB^{-1} = I$ 。一般长方形矩阵没有通常意义下的双侧逆矩阵。



向量值函数与 Jacobian 矩阵

若 $F: \mathbb{R}^n \rightarrow \mathbb{R}^m$, 则一阶导数可组织成

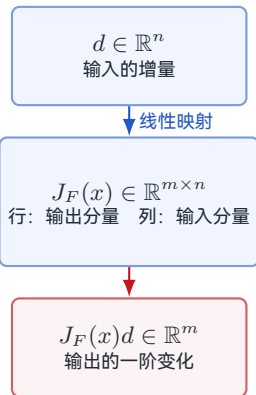
$$J_F(x) = \left[\frac{\partial F_i}{\partial x_j}(x) \right]_{i,j} \in \mathbb{R}^{m \times n}.$$

可微时,

$$F(x+d) \approx F(x) + J_F(x)d.$$

线性函数 $F(x) = Ax$ 的 Jacobian 就是 A 。

实值函数是 $m = 1$ 的情形, 此时 $J_f(x) = \nabla f(x)^\top$ 。
行对应输出分量, 列对应输入分量。



偏导数存在，不代表连续

定义

$$g(x, y) = \begin{cases} \frac{xy}{x^2 + y^2}, & (x, y) \neq (0, 0), \\ 0, & (x, y) = (0, 0). \end{cases}$$

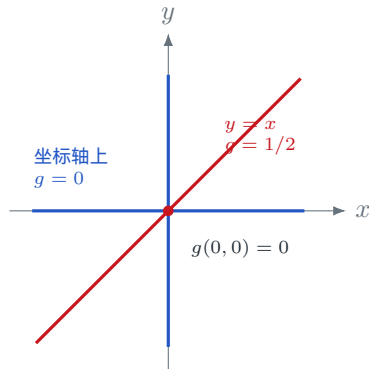
沿坐标轴 $g(h, 0) = g(0, h) = 0$, 所以

$$g_x(0, 0) = g_y(0, 0) = 0.$$

但沿 $y = x \neq 0$,

$$g(x, x) = \frac{1}{2} \neq g(0, 0).$$

所以它在原点不连续，更不可微。



线性代数复习

$$\text{设 } A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}, \quad \theta = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad y = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

1. 计算 $A\theta$ 、 $r = A\theta - y$ 、 $\|r\|_2$ 与 $\|r\|_2^2$ 。
2. 计算 $A^\top A$ ，说明 A 是否满列秩、 $\ker A$ 是否只有零向量。
3. 令 $u = (1, 2)^\top$ 、 $v = (2, -1)^\top$ ，计算 $u^\top v$ 与 $uv^\top$ 。

每一步先写出结果应有的维数，再进行计算。

$$\begin{matrix} A & (3 \times 2) \\ \theta & (2 \times 1) \end{matrix}$$



$$A\theta \quad (3 \times 1)$$



$$r = A\theta - y \quad (3 \times 1)$$

$$\|r\|_2^2 \quad (1 \times 1)$$

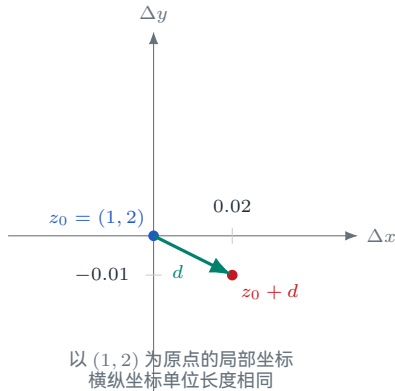
微分与链式法则复习

仍设 $f(x, y) = x^2 + xy + y^2$ 。

1. 求 $\nabla f(x, y)$ 及 $\nabla f(1, 2)$ 。
2. 用一阶近似估计 $f(1.02, 1.99)$ ，再与精确值比较。
3. 设 $x(t) = t$ 、 $y(t) = t^2$ ，分别用“先代入”和链式法则计算

$$\frac{d}{dt}f(t, t^2).$$

重点区分：偏导数、输入增量、实际函数增量和一阶近似。



二阶信息复习

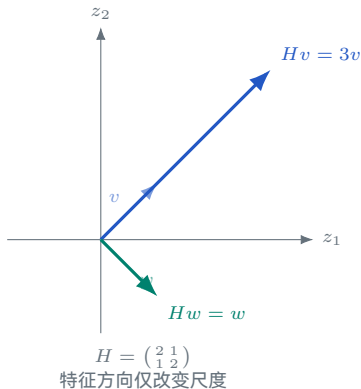
1. 求 $f(x, y) = x^2 + xy + y^2$ 的 Hessian, 并将 $f(x, y)$ 写成 $\frac{1}{2}z^\top H z$ 。

2. 判断下面的对称矩阵是否正定、是否半正定:

$$Q = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}.$$

3. 对 $H = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$, 验证 $(1, 1)^\top$ 与 $(1, -1)^\top$ 是特征向量, 并写出对应特征值。

优先用展开、配方和定义解释, 不必计算复杂特征多项式。





上海交通大学

SHANGHAI JIAO TONG UNIVERSITY

自然科学研究院

INSTITUTE OF NATURAL SCIENCES

Fin

