



維基百科

自由的百科全書

# 矩阵微积分

在数学中，**矩阵微积分**是多元微积分的一种特殊表达，尤其是在矩阵空间上进行讨论的时候。它把单个函数对多个变量或者多元函数对单个变量的偏导数写成向量和矩阵的形式，使其可以被当成一个整体被处理。这使得要在多元函数寻找最大或最小值，又或是要为微分方程系统寻解的过程大幅简化。这里我们主要使用统计学和工程学中的惯用记法，而张量下标记法更常用于物理学中。

## 记法

在本小节中，我们在表示向量和矩阵时，通过用单个变量来表示许多变量的方式，把矩阵记法的效用发挥到最大。接下来我们用不同字体来区分标量、向量和矩阵。我们使用 $M(n,m)$ 来表示包含 $n$ 行 $m$ 列的 $n\times m$ 实矩阵的空间。该空间中的一般矩阵用粗体大写字母表示，例如 $\mathbf{A}$ ， $\mathbf{X}$ ， $\mathbf{Y}$ 等。而若该矩阵属于 $M(n,1)$ ，即列向量，则用粗体小写字母表示，如 $\mathbf{a}$ ， $\mathbf{x}$ ， $\mathbf{y}$ 等。特别地， $M(1,1)$ 中的元素为标量，用小写斜体字母表示，如 $a$ ， $t$ ， $x$ 等。 $\mathbf{X}^T$ 表示矩阵转置， $\text{tr}(\mathbf{X})$ 表示矩阵的迹，而 $\det(\mathbf{X})$ 或 $|\mathbf{X}|$ 表示行列式。除非专门注明，所有函数都默认属于光滑函数 $C^1$ 。通常字母表前半部分的字母(a, b, c, ...)用于表示常量，而后半部分的字母(t, x, y, ...)用于表示变量。

## 向量求导

由于向量可看成仅有一列的矩阵，最简单的矩阵求导为向量求导。

这里的标记方法可以通过如下方式表达大部分向量微积分：把 $n$ 维向量构成的空间 $M(n,1)$ 等同为欧氏空间 $\mathbf{R}^n$ ，标量 $M(1,1)$ 等同于 $\mathbf{R}$ 。对应的向量微积分的概念在每小节末尾列出。

### 向量对标量求导

向量 $\mathbf{y} = [y_1 \ y_2 \ \cdots \ y_m]^T$ 关于标量 $x$ 的导数可以（用分子记法）写成

$$\frac{\partial \mathbf{y}}{\partial x} = \begin{bmatrix} \frac{\partial y_1}{\partial x} \\ \frac{\partial y_2}{\partial x} \\ \vdots \\ \frac{\partial y_m}{\partial x} \end{bmatrix}$$

这里的“分子布局记法”一般指，在表示导数向量（或矩阵）时，该导数的行数等于导数表达式 $\frac{\partial \mathbf{y}}{\partial x}$ 中处于分子部分的参数维度，即 $\mathbf{y}$ 的维度大小 $m$ ；若采用分母布局记法，则导数的行数等于导数表达式中处于分母部分的参数维度，即 $x$ 的维度大小。分子布局记法的结果与分母布局记法的结果互为转置关系。

在向量微积分中，向量 $\mathbf{y}$ 关于标量 $x$ 的导数也被称为向量 $\mathbf{y}$ 的切向量， $\frac{\partial \mathbf{y}}{\partial x}$ 。注意这里 $\mathbf{y} : \mathbb{R} \rightarrow \mathbb{R}^n$ 。

**例子** 简单的样例包括欧式空间中的速度向量，它是位移向量（看作关于时间的函数）的切向量。更进一步而言，加速度是速度的切向量。

## 标量对向量求导

标量 $y$ 对向量 $\mathbf{x} = [x_1 \ x_2 \ \cdots \ x_n]^T$ 的导数可以（用分子记法）写成

$$\frac{\partial y}{\partial \mathbf{x}} = \left[ \frac{\partial y}{\partial x_1} \quad \frac{\partial y}{\partial x_2} \quad \cdots \quad \frac{\partial y}{\partial x_n} \right]$$

在向量微积分中，标量 $y$ 在的空间 $\mathbb{R}^n$  (其独立坐标是 $\mathbf{x}$ 的分量)中的梯度是标量 $y$ 对向量 $\mathbf{x}$ 的导数的转置。在物理学中，电场是电势的负梯度向量。

标量函数 $f(\mathbf{x})$ 对空间向量 $\mathbf{x}$ 在单位向量 $\mathbf{u}$ （在这里表示为列向量）方向上的方向导数可以用梯度定义：

$$\nabla_{\mathbf{u}} f(\mathbf{x}) = \nabla f(\mathbf{x}) \cdot \mathbf{u}$$

使用刚才定义的标量对向量的导数的记法，我们可以把方向导数写作  $\nabla_{\mathbf{u}} f = \left( \frac{\partial f}{\partial \mathbf{x}} \right)^T \mathbf{u}$  这类记法在证明乘法法则和链式法则的时候非常直观，因为它们与我们熟悉的标量导数的形式较为相似。

## 向量对向量求导

前面两种情况可以看作是向量对向量求导在其中一个是一维向量情况下的特例。类似地我们将会发现有关矩阵的求导可被以一种类似的方式化归为向量求导。

向量函数 (分量为函数的向量)  $\mathbf{y} = [y_1 \ y_2 \ \cdots \ y_m]^T$  对输入向量 $\mathbf{x} = [x_1 \ x_2 \ \cdots \ x_n]^T$ 的导数，可以（用分子记法）写作

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_m}{\partial x_1} & \frac{\partial y_m}{\partial x_2} & \cdots & \frac{\partial y_m}{\partial x_n} \end{bmatrix}$$

在向量微积分中，向量函数 $\mathbf{y}$ 对分量表示一个空间的向量 $\mathbf{x}$ 的导数也被称为前推 (微分)，或雅可比矩阵。

向量函数 $\mathbf{f}$ 对 $\mathbb{R}^n$ 空间中向量 $\mathbf{v}$ 的前推为  $d\mathbf{f}(\mathbf{v}) = \frac{\partial \mathbf{f}}{\partial \mathbf{v}} d\mathbf{v}$

# 矩阵求导

有两种类型的矩阵求导可以被写成相同大小的矩阵：矩阵对标量求导和标量对矩阵求导。它们在解决应用数学的许多领域常见的最小化问题中十分有用。类比于向量求导，相应的概念有切矩阵和梯度矩阵。

## 矩阵对标量求导

矩阵函数 $\mathbf{Y}$ 对标量 $x$ 的导数被称为切矩阵，(用分子记法)可写成：

$$\frac{\partial \mathbf{Y}}{\partial x} = \begin{bmatrix} \frac{\partial y_{11}}{\partial x} & \frac{\partial y_{12}}{\partial x} & \cdots & \frac{\partial y_{1n}}{\partial x} \\ \frac{\partial y_{21}}{\partial x} & \frac{\partial y_{22}}{\partial x} & \cdots & \frac{\partial y_{2n}}{\partial x} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_{m1}}{\partial x} & \frac{\partial y_{m2}}{\partial x} & \cdots & \frac{\partial y_{mn}}{\partial x} \end{bmatrix}$$

## 标量对矩阵求导

定义在元素是独立变量的 $p \times q$ 矩阵 $\mathbf{X}$ 上的标量函数 $y$ 对 $\mathbf{X}$ 的导数可以(用分子记法)写作

$$\frac{\partial y}{\partial \mathbf{X}} = \begin{bmatrix} \frac{\partial y}{\partial x_{11}} & \frac{\partial y}{\partial x_{21}} & \cdots & \frac{\partial y}{\partial x_{p1}} \\ \frac{\partial y}{\partial x_{12}} & \frac{\partial y}{\partial x_{22}} & \cdots & \frac{\partial y}{\partial x_{p2}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y}{\partial x_{1q}} & \frac{\partial y}{\partial x_{2q}} & \cdots & \frac{\partial y}{\partial x_{pq}} \end{bmatrix}$$

定义矩阵上的重要的标量函数包括矩阵的迹和行列式。

类比于向量微积分，这个导数常被写成如下形式：

$$\nabla_{\mathbf{X}} y(\mathbf{X}) = \frac{\partial y(\mathbf{X})}{\partial \mathbf{X}}$$

类似地，标量函数 $f(\mathbf{X})$ 关于矩阵 $\mathbf{X}$ 在方向 $\mathbf{Y}$ 的方向导数可写成

$$\nabla_{\mathbf{Y}} f = \text{tr} \left( \frac{\partial f}{\partial \mathbf{X}} \mathbf{Y} \right)$$

梯度矩阵经常被应用在估计理论的最小化问题中，比如卡尔曼滤波算法的推导，因此在这些领域中有着重要的地位。

## 參考文獻

---

- Fang, Kai-Tai; Zhang, Yao-Ting. Generalized multivariate analysis. Science Press (Beijing) and Springer-Verlag (Berlin). 1990. [ISBN 3540176519](#). 9783540176510.
- Kollo, Tõnu; von Rosen, Dietrich. [Advanced multivariate statistics with matrices](#). Dordrecht: Springer. 2005. [ISBN 978-1-4020-3418-3](#).
- Pan, Jianxin; Fang, Kaitai. Growth curve models and statistical diagnostics. Beijing: Science Press. 2007. [ISBN 9780387950532](#).

## 延伸閱讀

---

- Lax, Peter D. 9. Calculus of Vector- and Matrix-Valued Functions. Linear algebra and its applications 2nd. Hoboken, N.J.: Wiley-Interscience. 2007. [ISBN 978-0-471-75156-4](#).

## 外部連結

---

检索自“<https://zh.wikipedia.org/w/index.php?title=矩阵微积分&oldid=80032926>”