雅可比矩阵与行列式

wikipedia

6 April 2022

在向量微积分中，多个变量的向量值函数的**雅可比矩阵** (**Jacobian matrix**) 是其所有一阶偏导数的矩阵。当这个矩阵是方阵时，也就是说，当函数的输入变量数与其输出向量分量数相同时，它的行列式被称为**雅可比行列式** (**Jacobian determinant**)。矩阵和 (如果适用的) 行列式在文献中通常简称为 “**Jacobian**”。[4]

假设 $\mathbf{f}:\mathbf{R}^n\to\mathbf{R}^m$ 是一个函数，使得它的每个一阶偏导数都存在于 $\mathbf{R}^n$ 上。该函数将点 $\mathbf{x}\in\mathbf{R}^n$ 作为输入，并生成向量 $\mathbf{f}(\mathbf{x})\in\mathbf{R}^m$ 作为输出。那么 $\mathbf{f}$ 的雅可比矩阵被定义为一个 $m\times n$ 矩阵，记为 $\mathbf{J}$，其第 (i,j) 项为 $\mathbf{J}_{ij}=\frac{\partial f_i}{\partial x_j}$，或明确地

$$\mathbf{J}=\begin{bmatrix}\frac{\partial \mathbf{f}}{\partial x_1} & \cdots & \frac{\partial \mathbf{f}}{\partial x_n}\end{bmatrix}=\begin{bmatrix}\nabla^{\mathrm{T}} f_1\\ \vdots \\ \nabla^{\mathrm{T}} f_m\end{bmatrix}=\begin{bmatrix}\frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n}\\ \vdots & \ddots & \vdots\\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n}\end{bmatrix}$$

其中 $\nabla^{\mathrm{T}} f_i$ 是梯度的转置 (行向量) 的第 i 个分量。

雅可比矩阵，其条目是 $\mathbf{x}$ 的函数，以各种方式表示；常用符号包括 $D\mathbf{f}$, $\mathbf{J}_{\mathbf{f}}$, $\nabla\mathbf{f}$, 以及 $\frac{\partial(f_1,\ldots,f_m)}{\partial(x_1,\ldots,x_n)}$。一些作者将雅可比定义为上述形式的转置。

雅可比矩阵表示 $\mathbf{f}$ 在 $\mathbf{f}$ 的可微分的每个点上的微分。具体来说，如果 $\mathbf{f}(\mathbf{x})$ 在 $\mathbf{x}$ 处可微[1]，如果 $\mathbf{h}$ 是一个列矩阵表示的位移向量，那么矩阵乘积 $\mathbf{J}(\mathbf{x})\cdot\mathbf{h}$ 是另一个位移向量，即 $\mathbf{f}$ 在 $\mathbf{x}$ 的邻域内的变化的最佳线性近似。这意味着将 $\mathbf{y}$ 映射到 $\mathbf{f}(\mathbf{x})+\mathbf{J}(\mathbf{x})\cdot(\mathbf{y}-\mathbf{x})$ 的函数是 $\mathbf{f}(\mathbf{y})$ 对于所有靠近 $\mathbf{x}$ 的点的最佳线性逼近。该线性函数称为 $\mathbf{f}$ 在 $\mathbf{x}$ 处的**导数** (*derivative*) 或**微分** (*differential*)。

当 $m=n$ 时，雅可比矩阵是方阵，因此其行列式是 $\mathbf{x}$ 的明确定义的函数，称为 $\mathbf{f}$ 的**雅可比行列式**。它携带有关 $\mathbf{f}$ 的局部行为的重要信息。特别是，函数 $\mathbf{f}$ 在点 $\mathbf{x}$ 的附近具有局部的反函数，当且仅当雅可比行列式在 $\mathbf{x}$ 处非零时，该逆函数是可微的 (参见雅可比猜想)。当改变多重积分中的变量时，雅可比行列式也会出现 (参见多个变量的替换规则)。

当 $m=1$ 时，即当 $f:\mathbf{R}^n\to\mathbf{R}$ 是标量值函数时，雅可比矩阵归约为行向量 $\nabla^{\mathrm{T}} f$；f 的所有一阶偏导数的这个行向量是 f 的梯度的转置，即 $\mathbf{J}_f=\nabla^T f$。进一步特化，当 $m=n=1$ 时，即当 $f:\mathbf{R}\to\mathbf{R}f$ 是单个变量的标量值函数时，雅可比矩阵只有一个条目；此条目是函数 f 的导数。

这些概念以数学家 Carl Gustav Jacob Jacobi (1804–1851) 的名字命名。

[1] $\mathbf{x}$ 的可微性意味着 $\mathbf{x}$ 上所有一阶偏导数的存在，但反过来，$\mathbf{x}$ 上所有一阶偏导数的存在并不意味着 $\mathbf{x}$ 的可微性，因此这是一个更强的条件。

目录

1 雅可比矩阵

多变量向量值函数的雅可比推广了多变量的标量值函数的梯度，进而推广了单个变量的标量函数的导数。换句话说，多个变量的标量值函数的雅可比矩阵是其梯度 (的转置)，单个变量的标量值函数的梯度是其导数。

在函数可微分的每个点，它的雅可比矩阵也可以被认为是描述函数在该点附近局部施加的“拉伸”、“旋转”或“变换”的量。例如，如果 $(x', y') = \mathbf{f}(x, y)$ 用于平滑变换图像，雅可比矩阵 $\mathbf{J}_{\mathbf{f}}(x, y)$ 描述了图像在 (x, y) 的邻域中如何被变换。

如果一个函数在一点上可微，它的微分由雅可比矩阵在坐标中给出。然而，一个函数不需要是可微的来定义它的雅可比矩阵，因为只需要存在它的一阶偏导数。

如果 $\mathbf{f}$ 在 $\mathbf{R}^n$ 中的点 $\mathbf{p}$ 处可微分，则其微分由 $\mathbf{J}_{\mathbf{f}}(\mathbf{p})$ 表示。在这种情况下，由 $\mathbf{J}_{\mathbf{f}}(\mathbf{p})$ 表示的线性变换是 $\mathbf{f}$ 在点 $\mathbf{p}$ 附近的最佳线性逼近，在这个意义上

$$\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{p}) = \mathbf{J}_{\mathbf{f}}(\mathbf{p})(\mathbf{x} - \mathbf{p}) + o(\|\mathbf{x} - \mathbf{p}\|) \quad (\text{ as } \mathbf{x} \to \mathbf{p}),$$

其中 $O(\|\mathbf{x}-\mathbf{p}\|)$ 是一个接近零的数量，其比 $\mathbf{x}$ 接近 $\mathbf{p}$ 时 $\mathbf{x}$ 和 $\mathbf{p}$ 之间的距离收敛快得多。该近似专门用于通过其一阶泰勒多项式来逼近单个变量的标量函数，即

$$f(x)-f(p)=f'(p)(x-p)+o(x-p) \quad (\text{ as } x \to p).$$

从这个意义上说，雅可比矩阵可以看作是多变量向量值函数的一种“一阶导数”。特别是，这意味着几个变量的标量值函数的梯度也可以被视为它的“一阶导数”。

可组合可微函数 $\mathbf{f}: \mathbf{R}^n \to \mathbf{R}^m$ 和 $\mathbf{g}: \mathbf{R}^m \to \mathbf{R}^k$ 满足链式法则，即对于 $\mathbf{R}^n$ 中的 $\mathbf{x}$，$\mathbf{J}_{\mathbf{g}\circ\mathbf{f}}(\mathbf{x})=\mathbf{J}_{\mathbf{g}}(\mathbf{f}(\mathbf{x}))\mathbf{J}_{\mathbf{f}}(\mathbf{x})$。

多个变量的标量函数的梯度雅可比矩阵有一个特殊的名称：Hessian matrix ，在某种意义上它是所讨论函数的“二阶导数”。

2 雅可比行列式

如果 $m=n$，则 $\mathbf{f}$ 是从 $\mathbf{R}^n$ 到自身的函数，雅可比矩阵是方阵。然后我们可以形成它的行列式，称为**雅可比行列式**。雅可比行列式有时简称 “the Jacobian”。

给定点的雅可比行列式给出了关于 $\mathbf{f}$ 在该点附近的行为的重要信息。例如，如果 $\mathbf{p}$ 处的雅可比行列式不为零，则连续可微函数 $\mathbf{f}$ 在点 $\mathbf{p} \in \mathbf{R}^n$ 附近可逆。这就是反函数定理。此外，如果 $\mathbf{p}$ 处的雅可比行列式为正数，则 $\mathbf{f}$ 保持 $\mathbf{p}$ 附近的方向；如果它是负数，$\mathbf{f}$ 反转方向。在 $\mathbf{p}$ 处的雅可比行列式绝对值给出了函数 $\mathbf{f}$ 在 $\mathbf{p}$ 附近扩大或缩小体积的因子；这就是它出现在一般替换规则中的原因。

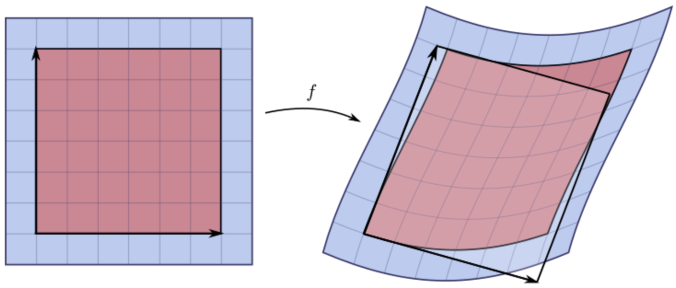

图 1: 非线性映射 $f: \mathbb{R}^2 \to \mathbb{R}^2$ 将一个小正方形 (左，红色) 发送到一个扭曲的平行四边形 (右，红色)。一个点的雅可比给出了该点附近的扭曲平行四边形的最佳线性近似 (右，半透明白色)，雅可比行列式给出了近似平行四边形面积与原始正方形面积之比。

雅可比行列式用于在评估函数在其域内的区域上的多重积分时更改变量。为了适应坐标的变化，雅可比行列式的大小作为积分内的乘法因子出现。这是因为 n 维 dV 元素在新坐标系中通常是平行六面体，而平行六面体的 n 维体积是其边向量的行列式。

雅可比行列也可用于通过近似平衡点附近的行为来确定微分方程组的平衡稳定性。其应用包括确定疾病建模中无病平衡的稳定性。[5]

3 逆矩阵

根据反函数定理，可逆函数的雅可比矩阵的逆矩阵就是逆函数的雅可比矩阵。也就是说，如果函数 $\mathbf{f}: \mathbf{R}^n \to \mathbf{R}^n$ 的雅可比在 $\mathbf{R}^n$ 中的点 $\mathbf{p}$ 处是连续且非奇异的，则当限制在 $\mathbf{p}$ 的某个邻域时 $\mathbf{f}$ 是可逆的，并且

$$\mathbf{J}_{\mathbf{f}^{-1}}=\mathbf{J}_{\mathbf{f}}^{-1}.$$

反之，如果雅可比行列式在某点不为零，则函数在该点附近**局部可逆** (*locally invertible*)，即在该点附近存在函数可逆的邻域。

在多项式函数的情况下，(未经证明的)雅可比猜想与全局可逆性有关，即由 n 个变量中的 n 个多项式定义的函数。它断言，如果雅可比行列式是一个非零常数 (或者，等价地，它没有任何复数零)，那么该函数是可逆的，并且它的逆是一个多项式函数。

4 临界点

如果 $\mathbf{f}: \mathbf{R}^n \to \mathbf{R}^m$ 是可微函数，则 $\mathbf{f}$ 的**临界点** (*critical point*) 是雅可比矩阵的秩不是最大的点。这意味着临界点的秩低于某个相邻点的秩。换句话说，设 k 是 $\mathbf{f}$ 的图像中包含的开球的最大维数；如果 $\mathbf{f}$ 的所有秩为 k 的所有子项都为零，则这个点是临界点。

在 $m = n = k$ 的情况下，如果雅可比行列式为零，则这个点是临界点。

5 例子

5.1 示例 1

考虑函数 $\mathbf{f}: \mathbf{R}^2 \to \mathbf{R}^2$，其中 $(x, y) \mapsto (f_1(x, y), f_2(x, y))$，由下式给出

$$\mathbf{f}\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} f_1(x, y) \\ f_2(x, y) \end{bmatrix} = \begin{bmatrix} x^2 y \\ 5x + \sin y \end{bmatrix}.$$

则我们有

$$f_1(x, y) = x^2 y$$

并且

$$f_2(x, y) = 5x + \sin y$$

并且 $\mathbf{f}$ 的雅可比矩阵是

$$\mathbf{J_f}(x, y) = \begin{bmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{bmatrix} = \begin{bmatrix} 2xy & x^2 \\ 5 & \cos y \end{bmatrix}$$

并且其雅可比行列式是

$$\det\left(\mathbf{J_f}(x, y)\right) = 2xy\cos y - 5x^2.$$

5.2 示例 2：极坐标-笛卡尔变换

从极坐标 (r, φ) 到笛卡尔坐标 (x, y) 的变换由函数 $\mathbf{F}: \mathbf{R}^+ \times [0, 2\pi) \to \mathbf{R}^2$ 给出，其分量为：

$$x = r\cos\varphi;$$

$$y = r\sin\varphi$$

$$\mathbf{J_F}(r, \varphi) = \begin{bmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \varphi} \end{bmatrix} = \begin{bmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{bmatrix}$$

雅可比行列式等于 r。这可用于转换两个坐标系之间的积分：

$$\iint_{\mathbf{F}(A)} f(x, y)dxdy = \iint_A f(r\cos\varphi, r\sin\varphi)rdrd\varphi.$$

5.3 示例 3：球面-笛卡尔变换

从球坐标 (ρ,φ,θ) [6] 到笛卡尔坐标 (x,y,z) 的变换由函数 $\mathbf{F}:\mathbf{R}^+\times[0,\pi)\times[0,2\pi)\to\mathbf{R}^3$ 给出，其分量为：

$$\begin{aligned}x&=\rho\sin\varphi\cos\theta;\\y&=\rho\sin\varphi\sin\theta;\\z&=\rho\cos\varphi.\end{aligned}$$

这种坐标变化的雅可比矩阵是

$$\mathbf{J}_{\mathbf{F}}(\rho,\varphi,\theta)=\begin{bmatrix}\frac{\partial x}{\partial\rho}&\frac{\partial x}{\partial\varphi}&\frac{\partial x}{\partial\theta}\\\frac{\partial y}{\partial\rho}&\frac{\partial y}{\partial\varphi}&\frac{\partial y}{\partial\theta}\\\frac{\partial z}{\partial\rho}&\frac{\partial z}{\partial\varphi}&\frac{\partial z}{\partial\theta}\end{bmatrix}=\begin{bmatrix}\sin\varphi\cos\theta&\rho\cos\varphi\cos\theta&-\rho\sin\varphi\sin\theta\\\sin\varphi\sin\theta&\rho\cos\varphi\sin\theta&\rho\sin\varphi\cos\theta\\\cos\varphi&-\rho\sin\varphi&0\end{bmatrix}.$$

行列式是 $\rho^2\sin\varphi$。由于 $dV=dxdydz$ 是矩形微分体积元的体积 (因为矩形棱柱的体积是其边的乘积)，我们可以将 $dV=\rho^2\sin\varphi d\rho d\varphi d\theta$ 解释为球面微分体积元素的体积。与矩形微分体积元的体积不同，这种微分体积元的体积不是常数，而是跟随坐标 (ρ 和 φ) 的变量。它可用于转换两个坐标系之间的积分：

$$\iiint_{\mathbf{F}(U)}f(x,y,z)dxdydz=\iiint_U f(\rho\sin\varphi\cos\theta,\rho\sin\varphi\sin\theta,\rho\cos\varphi)\rho^2\sin\varphi d\rho d\varphi d\theta.$$

5.4 示例 4

函数 $\mathbf{F}:\mathbf{R}^3\to\mathbf{R}^4$ 的分量如下：

$$\begin{aligned}y_1&=x_1\\y_2&=5x_3\\y_3&=4x_2^2-2x_3\\y_4&=x_3\sin x_1\end{aligned}$$

它的雅可比矩阵是

$$\mathbf{J}_{\mathbf{F}}(x_1,x_2,x_3)=\begin{bmatrix}\frac{\partial y_1}{\partial x_1}&\frac{\partial y_1}{\partial x_2}&\frac{\partial y_1}{\partial x_3}\\\frac{\partial y_2}{\partial x_1}&\frac{\partial y_2}{\partial x_2}&\frac{\partial y_2}{\partial x_3}\\\frac{\partial y_3}{\partial x_1}&\frac{\partial y_3}{\partial x_2}&\frac{\partial y_3}{\partial x_3}\\\frac{\partial y_4}{\partial x_1}&\frac{\partial y_4}{\partial x_2}&\frac{\partial y_4}{\partial x_3}\end{bmatrix}=\begin{bmatrix}1&0&0\\0&0&5\\0&8x_2&-2\\x_3\cos x_1&0&\sin x_1\end{bmatrix}.$$

这个例子表明雅可比矩阵不必是方阵。

5.5 示例 5

函数 $\mathbf{F}:\mathbf{R}^3\to\mathbf{R}^3$ 的分量如下：

$$\begin{aligned}y_1&=5x_2\\y_2&=4x_1^2-2\sin(x_2x_3)\\y_3&=x_2x_3\end{aligned}$$

它的雅可比矩阵是

$$\begin{vmatrix} 0 & 5 & 0 \\ 8x_1 & -2x_3 \cos(x_2 x_3) & -2x_2 \cos(x_2 x_3) \\ 0 & x_3 & x_2 \end{vmatrix} = -8x_1 \begin{vmatrix} 5 & 0 \\ x_3 & x_2 \end{vmatrix} = -40x_1 x_2.$$

由此我们看到，在 x_1 和 x_2 具有相同符号的那些点附近，**F** 反转了方向；除了 $x_1 = 0$ 或 $x_2 = 0$ 附近的点外，该函数在任何地方都是局部可逆的。直观地说，如果从点 $(1, 2, 3)$ 周围的一个小对象开始并将 **F** 应用于该对象，将得到一个体积约为原始对象体积的 $40 \times 1 \times 2 = 80$ 倍的结果对象，其中方向反转。

6 其他用途

6.1 回归与最小二乘拟合

雅可比矩阵在统计回归和曲线拟合中用作线性化设计矩阵；参见非线性最小二乘法。

6.2 动力系统

考虑一个形式为 $\dot{\mathbf{x}} = F(\mathbf{x})$ 的动力系统，其中 $\dot{\mathbf{x}}$ 是 $\mathbf{x}$ 相对于演化参数的时间 t 的 (按分量) 导数，而 $F : \mathbb{R}^n \to \mathbb{R}^n$ 是可微的。如果 $F(\mathbf{x}_0) = 0$，则 $\mathbf{x}_0$ 是一个稳定点(也称为稳态)。根据 Hartman-Grobman 定理，系统在驻点附近的行为与 $\mathbf{J}_F(\mathbf{x}_0)$, 的特征值有关，即稳定点 F 处的雅可比行列式。[7] 具体来说，如果特征值的实部都为负，则系统在驻点附近稳定，如果任何特征值的实部为正，则该点不稳定。如果特征值的最大实部为零，则雅可比矩阵不允许评估稳定性。[8]

6.3 牛顿法

一个耦合非线性方程的方系统可以通过牛顿法迭代求解。该方法使用方程组的雅可比矩阵。

7 参看

- Center manifold
- Hessian matrix
- Pushforward (differential)

8 参考文献

1. "Jacobian - Definition of Jacobian in English by Oxford Dictionaries". Oxford Dictionaries - English. Archived from the original on 1 December 2017. Retrieved 2 May 2018.

2. "the definition of jacobian". Dictionary.com. Archived from the original on 1 December 2017. Retrieved 2 May 2018.

3. Team, Forvo. "Jacobian pronunciation: How to pronounce Jacobian in English". forvo.com. Retrieved 2 May 2018.

4. W., Weisstein, Eric. "Jacobian". mathworld.wolfram.com. Archived from the original on 3 November 2017. Retrieved 2 May 2018.

5. Smith? RJ (2015). "The Joys of the Jacobian". Chalkdust. 2: 10–17.

6. Joel Hass, Christopher Heil, and Maurice Weir. Thomas' Calculus Early Transcendentals, 14e. Pearson, 2018, p. 959.

7. Arrowsmith, D. K.; Place, C. M. (1992). "The Linearization Theorem". Dynamical Systems: Differential Equations, Maps, and Chaotic Behaviour. London: Chapman & Hall. pp. 77–81. ISBN 0-412-39080-9.

8. Hirsch, Morris; Smale, Stephen (1974). Differential Equations, Dynamical Systems and Linear Algebra. ISBN 0-12-349550-4.

9 进一步阅读

- Gandolfo, Giancarlo (1996). "Comparative Statics and the Correspondence Principle". Economic Dynamics (Third ed.). Berlin: Springer. pp. 305–330. ISBN 3-540-60988-1.
- Protter, Murray H.; Morrey, Charles B., Jr. (1985). "Transformations and Jacobians". Intermediate Calculus (Second ed.). New York: Springer. pp. 412–420. ISBN 0-387-96058-9.

10 外部链接

- "Jacobian", Encyclopedia of Mathematics, EMS Press, 2001 [1994]
- Mathworld A more technical explanation of Jacobians