MAU 23205 Lecture 8

John Stalker

Trinity College Dublin

30 September 2021

Second derivatives

If f is a differentiable function from $\mathbf{R}^n$ (column vectors) to $\mathbf{R}$ then its derivative is

$$f'(\mathbf{x}) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(\mathbf{x}) & \cdots & \frac{\partial f}{\partial x_n}(\mathbf{x}) \end{bmatrix}.$$

This is a function from column vectors to row vectors, so not on the list of things with a sensible matrix derivative. Instead one takes the derivative of $(f')^T$, a function from column vectors to column vectors. This gives the Hessian matrix

$$\left((f')^T\right)'(\mathbf{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(\mathbf{x}) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{x}) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{x}) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(\mathbf{x}) \end{bmatrix}.$$

If $(f')^T$ is continuously differentiable we say that f is twice continuously differentiable. In that case the Hessian is a symmetric matrix.

A converse (Poincaré Lemma)

Suppose $\mathbf{g}$ is a continuously differentiable function from $\mathbf{R}^n$ (column vectors) to $\mathbf{R}^n$ $\mathbf{R}^n$ (column vectors). Is there a twice continuously differentiable function f from $\mathbf{R}^n$ to $\mathbf{R}$ such that $f' = \mathbf{g}^T$? Definitely no unless

$$\mathbf{g}'(\mathbf{x}) = \begin{bmatrix} \frac{\partial g_1}{\partial x_1}(\mathbf{x}) & \cdots & \frac{\partial g_1}{\partial x_n}(\mathbf{x}) \\ \vdots & \ddots & \vdots \\ \frac{\partial g_n}{\partial x_1}(\mathbf{x}) & \cdots & \frac{\partial g_n}{\partial x_n}(\mathbf{x}) \end{bmatrix}$$

is symmetric. Yes, if $\mathbf{g}$ is continuously differentiable on all of $\mathbf{R}^n$ and $\mathbf{g}'$ is symmetric there or, more generally, on a convex subset of $\mathbf{R}^n$, e.g. a ball, but not necessarily on general subsets of $\mathbf{R}^n$.

Application to invariants

Consider an autonomous system $\mathbf{x}'(t) = \mathbf{F}(\mathbf{x}(t)) = (\mathbf{F} \circ \mathbf{x})(t)$ Let $\mathbf{g} = Q\mathbf{F}$, where Q is an orthogonal antisymmetric matrix, i.e. $Q^{-1} = Q^T = -Q$. Then

$$(\mathbf{g}')^T = \mathbf{g}' \Leftrightarrow (\mathbf{F}')^T Q^T = Q\mathbf{F}' \Leftrightarrow (\mathbf{F}')^T Q^{-1} = Q\mathbf{F}' \Leftrightarrow (\mathbf{F}')^T = Q\mathbf{F}'Q.$$

If this happens then there's an f such that $f' = \mathbf{g}^T$. Then

$$(f \circ \mathbf{x})' = (f' \circ \mathbf{x})\,\mathbf{x}' = \left(\mathbf{g}^T \circ \mathbf{x}\right)(\mathbf{F} \circ \mathbf{x}) = \left(\mathbf{g}^T\mathbf{F}\right) \circ \mathbf{x} = \left(\mathbf{F}^T Q^T \mathbf{F}\right) \circ \mathbf{x}$$

1×1 matrices are symmetric so

$$\mathbf{F}^T Q^T \mathbf{F} = \left(\mathbf{F}^T Q^T \mathbf{F}\right)^T = \mathbf{F}^T Q \mathbf{F} = -\mathbf{F}^T Q^T \mathbf{F}$$

Therefore $\mathbf{F}^T Q^T \mathbf{F} = 0$, $(f \circ \mathbf{x})' = 0$ and $f \circ \mathbf{x}$ is constant. In other words, f is an invariant.

The 2×2 case

$Q = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ is orthogonal and antisymmetric. I'll write $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ and $\mathbf{F} = \begin{bmatrix} F \\ G \end{bmatrix}$ instead of $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ and $\mathbf{F} = \begin{bmatrix} F_1 \\ F_2 \end{bmatrix}$ to avoid subscripts. The system is then

$$x'(t) = F(x(t), y(t)), \qquad y'(t) = G(x(t), y(t)).$$

$$\mathbf{F}' = \begin{bmatrix} \frac{\partial F}{\partial x} & \frac{\partial F}{\partial y} \\ \frac{\partial G}{\partial x} & \frac{\partial G}{\partial y} \end{bmatrix}.$$

The condition $(\mathbf{F}')^T = Q\mathbf{F}'Q$ is then $\frac{\partial F}{\partial x} = -\frac{\partial G}{\partial y}$. If this is satisfied then there is an invariant. It satisfies the equation $(f')^T = Q\mathbf{F}$, i.e. $\frac{\partial f}{\partial x} = -G$ and $\frac{\partial f}{\partial y} = F$.

Integrals of matrix valued functions of a real variable

It's possible to define the integral of a matrix valued function on an interval. This may or may not follow the construction of the integral of real valued functions. There are two main constructions, one of which generalises nicely and the other doesn't. In any case, it can be defined and has the expected properties.

This includes the Fundamental Theorem of Calculus. If Q is a continuous matrix valued function on an interval $[a, b]$ and P is defined by

$$P(x) = \int_{[a,x]} Q(y)\, dy$$

for $x \in [a, b]$ then $P' = Q$. If P and Q are matrix valued functions on an interval $[a, b]$ and $P' = Q$ and Q is integrable there then

$$\int_{[a,b]} Q(y)\, dy = P(b) - P(a).$$

The Poincaré Lemma again

As we saw, if $\mathbf{g}$ is continuously differentiable in a convex set in $\mathbf{R}^n$ and $\mathbf{g}'$ is symmetric there then there is an f such that $f' = \mathbf{g}^T$. Suppose $\mathbf{x}_*$ and $\mathbf{x}$ are points in this convex set. Let $\mathbf{y}(s) = (1-s)\mathbf{x}_* + s\mathbf{x}$. Then $\mathbf{y}'(s) = (\mathbf{x} - \mathbf{x}_*)$. Also $(f \circ \mathbf{y})' = (f' \circ \mathbf{y})\mathbf{y}' = (\mathbf{g}^T \circ \mathbf{y})(\mathbf{x} - \mathbf{x}_*)$

$$
\begin{aligned}
f(\mathbf{x}) - f(\mathbf{x}_*) &= (f \circ \mathbf{y})(1) - (f \circ \mathbf{y})(0) \\
&= \int_{[0,1]} (f \circ \mathbf{y})'(s)\, ds \\
&= \int_{[0,1]} (\mathbf{g}^T \circ \mathbf{y})(s)(\mathbf{x} - \mathbf{x}_*)(s)\, ds
\end{aligned}
$$

$$
f(\mathbf{x}) = f(\mathbf{x}_*) + \int_{[0,1]} (\mathbf{g}^T \circ \mathbf{y})(s)(\mathbf{x} - \mathbf{x}_*)(s)\, ds
$$

This isn't a good way to compute f, but it is used to show that it exists. The convexity is needed to make sure $\mathbf{y}(s)$ stays with the set for $s \in [0,1]$.

The Inverse Function Theorem

Suppose $\mathbf{u}$ is a continuously differentiable function from a ball about $\mathbf{x}_* \in \mathbf{R}^n$ to $\mathbf{R}^n$ (column vectors in both cases). If $\mathbf{u}'(\mathbf{x}_*)$ is invertible then there's a function $\mathbf{v}$ from a ball about $\mathbf{u}(\mathbf{x}_*)$ to $\mathbf{R}^n$ such that $(\mathbf{u} \circ \mathbf{v})(\mathbf{y}) = \mathbf{y}$ for all $\mathbf{y}$ in an open ball about $\mathbf{u}(\mathbf{x}_*)$ and $(\mathbf{v} \circ \mathbf{u})(\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x}$ in a ball about $\mathbf{x}_*$. Also, $\mathbf{v}'(\mathbf{y}) = ((\mathbf{u}' \circ \mathbf{v})(\mathbf{y}))^{-1}$ in the ball where it's defined. If the balls are chosen small enough this $\mathbf{v}$ is unique.
This $\mathbf{v}$ is called the inverse function $\mathbf{u}$.
In the case $\mathbf{R}$ this is the familiar Inverse Function Theorem for real valued functions of a real variable.

Implicit Function Theorem

Suppose $\mathbf{G}$ is a continuously differentiable function from a ball about the point $(\mathbf{x}_*, \mathbf{y}_*)$ in $\mathbf{R}^m \times \mathbf{R}^n$ to $\mathbf{R}^n$. Let $\frac{\partial G}{\partial x}$ be the matrix consisting of the first m columns of $\mathbf{G}'$ and $\frac{\partial G}{\partial y}$ the matrix consisting of the last n columns of $\mathbf{G}'$. Suppose $\frac{\partial G}{\partial y}(\mathbf{x}_*, \mathbf{y}_*)$ is invertible. Then there is a function $\mathbf{F}$ from a ball about $\mathbf{x}_*$ in $\mathbf{R}^m$ to $\mathbf{R}^n$ such that $\mathbf{F}(\mathbf{x}_*) = \mathbf{y}_*$ and $\mathbf{G}(\mathbf{x}, \mathbf{F}(\mathbf{x})) = \mathbf{G}(\mathbf{x}_*, \mathbf{y}_*)$ for all $\mathbf{x}$ in the ball where $\mathbf{F}$ is defined. Also

$$\mathbf{F}'(x) = -\left(\frac{\partial \mathbf{G}}{\partial \mathbf{y}}(\mathbf{x}, \mathbf{F}(\mathbf{x}))\right)^{-1} \frac{\partial \mathbf{G}}{\partial \mathbf{x}}(\mathbf{x}, \mathbf{F}(\mathbf{x})).$$

If the balls are chosen small enough this $\mathbf{F}$ is unique.
This was used in Lecture 4 to convert a system of the form $\mathbf{G}(\mathbf{x}(t), \mathbf{x}'(t)) = \mathbf{0}$ into one of the form $\mathbf{x}'(t) = \mathbf{F}(\mathbf{x}(t))$.

Invariants, again

Suppose F and G are continuously differentiable in a ball about (x_0, y_0) and $\frac{\partial F}{\partial x} = -\frac{\partial G}{\partial y}$ there. Suppose also that $F(x_0, y_0) \neq 0$. As we saw, there's a function f from that ball to $\mathbf{R}$ such that $\frac{\partial f}{\partial x} = -G$ and $\frac{\partial f}{\partial y} = F$. Then $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$. By the implicit function theorem there is a function y defined in a ball about x_0, such that $y(x_0) = y_0$ and $f(x, y(x)) = f(x_0, y_0)$ in that ball. Differentiating $f(x, y(x)) = f(x_0, y_0)$,

$$\frac{\partial f}{\partial x}(x, y(x)) + \frac{\partial f}{\partial y}(x, y(x))y'(x) = 0.$$

$$-G(x, y(x)) + F(x, y(x))y'(x) = 0.$$

We've "solved" the initial value problem

$$y'(x) = \frac{G(x, y(x)}{F(x, y(x))} \qquad y(x_0) = y_0.$$

Equations of this form, with $\frac{\partial F}{\partial x} = -\frac{\partial G}{\partial y}$, are called *integrable*.