SC 639 (Spring 2020) - Mathematical Structures for Control

Ravi N. Banavar [1]
Ashutosh Jindal [1]

[1] Systems and Control Engineering,
IIT Bombay, India

July 30, 2020

Outline

Outline

Partial Derivatives

Definition

Let U be an open subset of $\mathbb{R}^n$, and let $f : U \mapsto \mathbb{R}^n$. We define partial derivative at a point $x \in U$ by

$$\begin{aligned} D_i f(x) &= \lim_{h \to 0} \frac{f(x + he_i) - f(x)}{h} \\ &= \lim_{h \to 0} \frac{f(x_1, \ldots, x_i + h, \ldots, x_n) - f(x_1, \ldots, x_n)}{h} \end{aligned}$$

If this limit exists, $D_i f(x)$ is the partial derivative with respect to x_i evaluated at x and is denoted by

$$D_i f(x) = \frac{\partial f}{\partial x_i}(x)$$

Outline

Differentiable Function

A function $f : U \to \mathbb{R}$ is said to be differentiable at a point $x \in U$ if there exist a vector $A \in \mathbb{R}^n$ and $h \in \mathbb{R}^n$ and a function $\psi : \mathbb{R}^n \to \mathbb{R}$ such that

$$f(x+h) = f(x) + \langle A, h \rangle + \|h\| \, \psi(h)$$

and ψ satisfies

$$\lim_{h \to 0} \psi(h) = 0$$

Gradient

Let $f(\cdot) : U \to \mathbb{R}$, and all partial derivatives $D_i f$ exist at $x \in U$ then the gradient is defined by

$$\nabla f(x) := \left(\frac{\partial f}{\partial x_1}(x), \ldots, \frac{\partial f}{\partial x_n}(x) \right)$$

Gradient and Differentiability

$f : U \to \mathbb{R}$ is differentiable at a point $x \in U$ if and only if $\nabla f(x)$ exists at x and

$$A = \nabla f(x)$$

Chain Rule

Let $\varphi : J \to \mathbb{R}^n$ be a differentiable function defined on some interval J and values in some open set $U \subset \mathbb{R}^n$. Let be $f : U \to \mathbb{R}^n$ a differentiable function, then $f \circ \varphi : J \to \mathbb{R}$ is differentiable and for a given $t \in J$

$$(f \circ \varphi)'(t) = \nabla f(\varphi(t)).\varphi'(t)$$

Gradient and Tangent Space

Directional Derivative

Let x be a point of U and let v be a fixed vector with $\|v\| = 1$. Directional derivative of f at x in the direction of v is given by

$$D_v f(x) = \frac{d}{dt} f(x + tv)\bigg|_{t=0} = \langle \nabla f(x), v \rangle$$

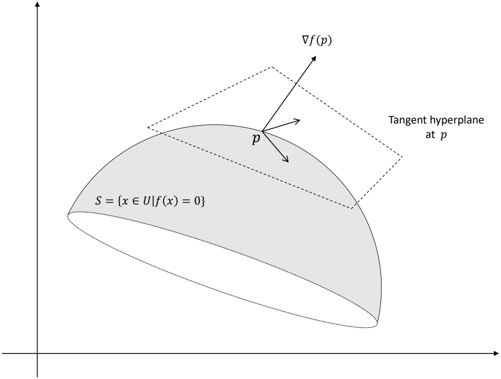
Direction of $\nabla f(x)$ is the direction of maximal increase of the function f at x

Hypersurface and Tangent Hyperplane

Let f be a differentiable function on some open set U in $\mathbb{R}^n$. Let $c \in \mathbb{R}$ and let S be the set of points x such that

$$f(x) = c, \quad \text{but} \quad \nabla f(x) \neq 0.$$

The set S is called a hypersurface in $\mathbb{R}^n$. The tangent hyperplane of S at a point $p \in S$ is defined as the hyperplane passing through p and perpendicular to $\nabla f(p)$.



Level Surface S and the gradient $\nabla f(p)$ at a point $p \in S$

Outline

Derivatives as a Linear Map

Space of Continuous Linear Maps

Let E, F be normed vector spaces. We denote the space of continuous linear maps $\lambda : E \to F$ by $L(E, F)$. $L(E, F)$ assumes a vector space structure. For $\lambda_1, \lambda_2 \in L(E, F)$ and $c \in \mathbb{R}$

$$(\lambda_1 + \lambda_2)(x) = \lambda_1(x) + \lambda_2(x)$$
$$(c\lambda)(x) = c\lambda(x)$$

Derivative as a Linear Map

Let U be open in E, and let $x \in U$. Let $f : U \to F$ be a map. The f is said to be differentiable at x if there exists a continuous linear map $\lambda : E \to F$ and a map ψ defined for all sufficiently small $h \in E$ and values in F, such that

$$\lim_{h \to 0} \psi(h) = 0$$

and

$$f(x + h) = f(x) + \lambda(h) + \|h\| \, \psi(h)$$

Jacobian Matrix

Let U be an open set of $\mathbb{R}^n$, and $f : U \to \mathbb{R}^m$ be a differentiable map at x. Then the continuous linear map is represented by the matrix

$$Df(x) = \frac{\partial f}{\partial x}(x)$$

and the matrix

$$Df(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \dots & \frac{\partial f_1}{\partial x_n}(x) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \dots & \frac{\partial f_m}{\partial x_n}(x) \end{pmatrix}$$

is called the Jacobian of f at x.

Properties of the Derivative

- **Sum** : Let $U \subset E$ be a open set and $f, g : U \to F$ be differentiable at $x \in U$. Then $f + g$ is differentiable at x and

$$(f+g)'(x) = f'(x) + g'(x)$$

and for $c \in \mathbb{R}$

$$(cf)'(x) = cf'(x)$$

- **Product** : Let $f, g : U \to F$ be differentiable at $x \in U$. Then the product map fg is differentiable at x, and

$$(fg)'(x) = f'(x)g(x) + f(x)g'(x)$$

- **Chain Rule** : Let U be open in E and V be open in F. Let $f : U \to V$ and $g : V \to G$ be functions differentiable at $x \in U$ and $f(x) \in V$ respectively. Then $g \circ f$ is differentiable at x and

$$(g \circ f)'(x) = g'(f(x)) \circ f'(x)$$

Second Derivative

Definition

Let U be open in E and let $f : U \to F$ be a differentiable map. Then if it exists, the second derivative is a map defined as

$$D^2 f = f^{(2)} : U \to L(E, L(E, F)) \tag{1}$$

Theorem

Let U be open in E and $f : U \to F$ be twice differentiable and such that $D^2 f$ is continuous. Then for each $x \in U$ the bilinear map $D^2 f$ is symmetric i.e.

$$D^2 f(x)(v, w) = D^2 f(x)(w, v)$$

Hessian of f

Theorem

Let U be open in $\mathbb{R}^n$ and let $f : U \to \mathbb{R}$ be a function. Then f is of class C^2 if and only if all partial derivatives of f upto order ≤ 2 exists and are continuous.

Hessian

The matrix representation of $D^2 f(x)$ is called the Hessian of f at x and is denoted by

$$\nabla^2 f(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x) & \dots & \frac{\partial^2 f}{\partial x_n^2}(x) \end{pmatrix}$$

Outline

Inverse Function Theprem

Theorem

Inverse Function Theorem *: Let U be open in E, let $x_0 \in U$ and let $f : U \to F$ be a C^1 map. Assume the derivative $f'(x_0) : E \to F$ is invertible. Then f is locally C^1-invertible at x_0. If φ is its local inverse, and $y = f(x)$, then $\varphi'(y) = f'(x)^{-1}$*

Lemma

Shrinking Lemma *: Let M be a closed subset of complete normed vector space. Let $f : M \to M$ be a mapping, and assume there exist a number k, $0 < k < 1$, such that for all $x, y \in M$ we have*

$$\|f(x) - f(y)\| \leq k \|x - y\|$$

Then f has a unique fixed point, i.e. there exists a unique $x_0 \in M$ such that $f(x_0) = x_0$.

Proof of Inverse Function Theorem

- **Assumption** : Without loss of generality, set $x_0 = 0$, $f(0) = 0$, and $f'(0) = I$.
- Let $g(x) = x - f(x)$ such that $g'(0) = 0$. By continuity there exists an $r > 0$ such that
$$\|g'(x)\| \leq \frac{1}{2} \quad \forall \ \|x\| \leq r.$$
- Continuity of f', and $f'(0) = 0$ implies $f'(x)$ is invertible for $\|x\| \leq r$
- From mean value theorem we have $\|g(x)\| \leq \frac{1}{2}\|x\|$ i.e. g maps closed ball $\bar{B}_r(0)$ into closed ball $\bar{B}_{r/2}(0)$
- Define $g_y(x) = y + x - f(x)$, so that it has a unique fixed point (guaranteed by shrinking lemma) at $f(x) = y$.
- Let $U_1 = \{x \in B_r(0) : \|f(x)\| < \boldsymbol{r}/2\}$ and $V_1 = f(U_1)$ be its image. Since $f : U_1 \to V_1$ is injective, inverse map exist
$$f : U_1 \to V_1 \quad f^{-1} = \varphi : V_1 \to U_1.$$

To show : V_1 is open φ is of class C^1

Proof (contd.)

- Let $x_1 \in U_1$ and let $y_1 = f(x_1)$ so that $\|y_1\| < r/2$.
- For $y \in E$ such that $\|y\| < r/2$ there exist a unique $x \in \bar{B}_r(0)$ such that $f(x) = y$. Then we have

$$\begin{aligned} \|x - x_1\| &\leq \|f(x) - f(x_1)\| + \|g(x) - g(x_1)\| \\ &\leq \|f(x) - f(x_1)\| + \frac{1}{2}\|x - x_1\| \end{aligned}$$

$$\|x - x_1\| \leq 2\|f(x) - f(x_1)\| \qquad (*)$$

- Hence, y is sufficiently close to y_1, if x is sufficiently close to x_1, thus $= f^{-1}$ is continuous.
- If $x \in U_1$, then $y \in V_1$ and hence V_1 is open.

Proof (contd.)

- To conclude we prove differentiability of $\varphi = f^{-1}$. We know $f'(x_1)$ is invertible because

$$f(x) - f(x_1) = f'(x_1)(x - x_1) + \|x - x_1\| \, \psi(x - x_1)$$

where $\lim_{x - x_1} \psi(x - x_1) = 0$.

- Substitute above result in

$$f^{-1}(y) - f^{-1}(y_1) - f'(x_1)^{-1}(y - y_1) = x - x_1 - f'(x_1)^{-1}(f(x) - f(x_1)) \quad (**)$$

- Using $(*)$ and a bound C for $f'(x_1)^{-1}$, we obtain

$$\|(**)\| \leq 2C \, \|y - y_1\| \, \|\psi(\varphi(y) - \varphi(y_1))\|$$

- Continuity of $\varphi = f^{-1}$ implies $\varphi'(y_1) = f'(x_1)^{-1}$. Thus, we have

$$\varphi'(y) = f'(\varphi(y))^{-1}$$

which is continuous. Thus φ is of class C^1, there by completing the proof.

- **Corrollary**: If f is of class C^p then its local inverse is of class C^p

Outline

Implicit Function Theorem

Theorem

Let $f : U_1 \times U_2 \to \mathbb{R}$ be a function of two real variables defined on a product of two open intervals U_1, U_2. Assume that f is of class C^p. Let $(a, b) \in U_1 \times U_2$ such that $f(a, b) = 0$ and $D_2 f(a, b) \neq 0$. Then the map

$$\psi : U_1 \to U_1 \times U_2 \to \mathbb{R} \times \mathbb{R}$$

given by

$$(x, y) \mapsto (x, f(x, y))$$

is locally C^p invertible at (a, b)

Proof.

The Jacobian matrix of ψ at (a, b) is given by

$$D\psi(x, y) = \begin{pmatrix} 1 & 0 \\ \frac{\partial f}{\partial x}(a, b) & \frac{\partial f}{\partial y}(a, b) \end{pmatrix}$$

is nonsingular at (a, b). The inverse mapping guarantees that ψ is locally invertible at (a, b). $\square$

Theorem

Implicit Function Theorem *: Let $f : U_1 \times U_2 \to \mathbb{R}$ be a function of two variables, defined on product of open interval. Let $(a, b) \in J_1 \times J_2$ such that $f(a, b) = 0$ and $D_2 f(a, b) \neq 0$. Then there exists an open interval $J \in \mathbb{R}$ containing a and a C^p function $g : J \to \mathbb{R}$ such that*

$$g(a) = b \quad \text{and } f(x, g(x)) = 0 \text{ for all } x \in J$$

Proof.

- $\psi : U_1 \times U_2 \to \mathbb{R} \times \mathbb{R}$ given by $(x, y) \mapsto (x, f(x, y))$ is locally invertible at (a, b).
- Let its local inverse be $\varphi = \psi^{-1} = (\varphi_1, \varphi_2)$ such that $\varphi(x, z) = (x, \varphi_2(x, z))$ and let $g(x) = \varphi_2(x, 0)$.
- $\psi(a, b) = (a, 0)$ implies $\varphi_2(a, 0) = b$ i.e. $g(a) = b$ Since ψ and φ are inverse mappings, we have

$$(x, 0) = \psi(\varphi(x, 0)) = \psi(x, g(x)) = (x, f(x, g(x)))$$

i.e. $f(x, g(x)) = 0$, proving the result. □

Theorem

Implicit function theorem (for $\mathbb{R}^n$): *Let U be open in $\mathbb{R}^n$ and $f : U \to \mathbb{R}$ be a C^p function on U. Let $(a, b) = (a_1, \ldots, a_{n-1}, b) \in U$ such that $f(a, b) = 0$ and $D_n f(a, b) \neq 0$. Then there exists an open ball V in $\mathbb{R}^{n-1}$ centered at a and a C^p function*

$$g : V \to \mathbb{R}$$

such that

$$g(a) = b \quad \text{and } f(x, g(x)) = 0 \text{ for all } x \in V$$

Proof.

The proof is similar to the implicit function theorem for two variables. □