Week 9 Recitation

Potential Function and Second Derivative

Yahoo

UM-SJTU Joint Institu(te

Summer 2017

Gradient

Definition

The transpose of the Jacobian is the gradient

$$\nabla f(x) := (J_f(x))^T = \begin{pmatrix} \frac{\partial f}{\partial x_1}\Big|_x \\ \vdots \\ \frac{\partial f}{\partial x_n}\Big|_x \end{pmatrix}$$

Gradient

Definition

The transpose of the Jacobian is the gradient

$$\nabla f(x) := (J_f(x))^T = \begin{pmatrix} \frac{\partial f}{\partial x_1}\Big|_x \\ \vdots \\ \frac{\partial f}{\partial x_n}\Big|_x \end{pmatrix}$$

Comment.

The gradient acts on a map from $\mathbb{R}^n$ to $\mathbb{R}$.

Directional Derivative

Definition

$h \in \mathbb{R}^n$, $\|h\| = 1$, the directional derivative is

$$D_h f|_x := \frac{d}{dt} f(x + th)\Big|_{t=0}.$$

Directional Derivative

Definition

$h \in \mathbb{R}^n$, $\|h\| = 1$, the directional derivative is

$$D_h f|_x := \frac{d}{dt} f(x + th)\Big|_{t=0}.$$

Question.

Calculate the directional derivative of function $f(x_1, x_2) = x_1^2 + x_2^2$

Directional Derivative

Question.

Calculate the directional derivative of function $f(x_1, x_2) = x_1^2 + x_2^2$

Solution.

$$\begin{aligned} D_h f\Big|_x &= \frac{d}{dt} f(x + th)\Big|_{t=0} \\ &= \frac{d}{dt}[(x_1 + th_1)^2 + (x_2 + th_2)^2]\Big|_{t=0} \\ &= 2x_1 h_1 + 2x_2 h_2 \end{aligned}$$

Directional Derivative

Comment.

$$D_h f\Big|_x = \langle \nabla f, h \rangle$$

Directional Derivative

Comment.

$$D_h f\Big|_x = \langle \nabla f, h \rangle$$

Solution.

$$\nabla f(x) = (\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}) = (2x_1, 2x_2)$$

$$\begin{aligned} D_h f\Big|_x &= (2x_1, 2x_2) \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \\ &= 2x_1 h_1 + 2x_2 h_2 \end{aligned}$$

Normal Derivative

Definition

Normal derivative is directional derivative along the normal vector direction of a curve

$$\left.\frac{\partial f}{\partial n}\right|_p := D_{N(p)} f|_p$$

Normal Derivative

Definition

Normal derivative is directional derivative along the normal vector direction of a curve

$$\left.\frac{\partial f}{\partial n}\right|_p := D_{N(p)} f|_p$$

Question.

Find the normal derivative of function $f(x_1, x_2) = x_1^2 + x_2^2$ on the curve $\{(x_1, x_2) \in \mathbb{R}^2 : x_2 = x_1^2\}$

Normal Derivative

Question.

Find the normal derivative of function $f(x_1, x_2) = x_1^2 + x_2^2$ on the curve $\{(x_1, x_2) \in \mathbb{R}^2 : x_2 = x_1^2\}$

Solution.

Let $\gamma(t) = (t, t^2)$. The normal vector of function $f(x_1, x_2) = x_1^2 + x_2^2$ on the curve $\{(x_1, x_2) \in \mathbb{R}^2 : x_2 = x_1^2\}$ is

$$N \circ \gamma(t) = \frac{(T \circ \gamma)'(t)}{\|(T \circ \gamma)'(t)\|} = \frac{1}{\sqrt{1 + 4t^2}} \begin{pmatrix} -2t \\ 1 \end{pmatrix}$$

At a point $p = \gamma(t)$ on $\mathcal{C}$ the normal derivative is hence

$$\left.\frac{\partial f}{\partial t}\right|_{\gamma(t)} = \langle \nabla f(\gamma(t)), N \circ \gamma(t) \rangle = \frac{-2t^2}{\sqrt{1 + 4t^2}}$$

Gradient

- $\nabla f(x)$ points in the direction of the greatest directional derivative of f at x.

$$D_h f\Big|_x = \langle \nabla f, h \rangle = |\nabla f(x)| \cos \angle \langle \nabla f, h \rangle$$

- $\nabla f(x)$ is perpendicular to the contour line of f at x.

Tangent Plane to the Graph of a Function

2.4.6. Definition. Let $\Omega \subset \mathbb{R}^n$ be open and $f\colon \Omega \to \mathbb{R}$ differentiable at $x_0 \in \Omega$. Then the equation

$$x_{n+1} = Tf(x; x_0), \qquad\qquad x = (x_1, \dots, x_n) \in \mathbb{R}^n,$$

defines the *tangent plane* to the graph $\Gamma(f) \in \mathbb{R}^n \times \mathbb{R}$ of f at the point $(x_0, f(x_0)) \in \mathbb{R}^{n+1}$.

Tangent Plane to the Graph of a Function

The tangent plane of a function $f\colon \mathbb{R}^2 \to \mathbb{R}$ at (x_0, y_0) is given by

$$
\begin{aligned}
z &= f(x_0, y_0) + \left(\frac{\partial f}{\partial x}(x_0, y_0),\ \frac{\partial f}{\partial y}(x_0, y_0)\right)\left(\begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}\right) \\
&= f(x_0, y_0) + (x - x_0)\frac{\partial f}{\partial x}(x_0, y_0) + (y - y_0)\frac{\partial f}{\partial y}(x_0, y_0)
\end{aligned}
$$

$$
\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \\ f(x_0, y_0) \end{pmatrix} + s\begin{pmatrix} 1 \\ 0 \\ \frac{\partial f}{\partial x}(x_0, y_0) \end{pmatrix} + t\begin{pmatrix} 0 \\ 1 \\ \frac{\partial f}{\partial y}(x_0, y_0) \end{pmatrix},
$$

which defines a plane in $\mathbb{R}^3$. The vectors

$$
t_1 := \begin{pmatrix} 1 \\ 0 \\ \frac{\partial f}{\partial x}(x_0, y_0) \end{pmatrix} \qquad \text{and} \qquad t_2 := \begin{pmatrix} 0 \\ 1 \\ \frac{\partial f}{\partial y}(x_0, y_0) \end{pmatrix}
$$

Normal Vector to the Graph of a Function

we can find a vector orthogonal to the tangent plane by taking the vector product

$$
\begin{aligned}
n &= \begin{pmatrix} 1 \\ 0 \\ \frac{\partial f}{\partial x} \end{pmatrix} \times \begin{pmatrix} 0 \\ 1 \\ \frac{\partial f}{\partial y} \end{pmatrix} = \det \begin{pmatrix} e_1 & e_2 & e_3 \\ 1 & 0 & \frac{\partial f}{\partial x} \\ 0 & 1 & \frac{\partial f}{\partial y} \end{pmatrix} \\
&= \begin{pmatrix} -\frac{\partial f}{\partial x} \\ -\frac{\partial f}{\partial y} \\ 1 \end{pmatrix}
\end{aligned}
$$

Tangent Plane to the Graph of a Function

Example

Find the tangent plane and normal line of the curve at point $(1, 2, 6)$

$$f(x, y) = 4x^2 + y$$

Tangent Plane to the Graph of a Function

Solution.

The partial derivative of f over x and y at $(1,2)$ is

$$\frac{\partial f}{\partial x}(1,2) = 8x\Big|_{x=1} = 8$$

$$\frac{\partial f}{\partial x}(1,2) = 1\Big|_{y=2} = 1$$

The tangent plane is $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 6 \end{pmatrix} + s\begin{pmatrix} 1 \\ 0 \\ 8 \end{pmatrix} + t\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$

Normal vector is $\begin{pmatrix} -\frac{\partial f}{\partial x}(1,2) \\ -\frac{\partial f}{\partial x}(1,2) \\ 1 \end{pmatrix} = \begin{pmatrix} -8 \\ -1 \\ 1 \end{pmatrix}$

The Second Derivative

2.5.1. Definition. Let X, V be finite-dimensional normed vector spaces and $\Omega \subset X$ an open set. A function $f : \Omega \to V$ is said to be *twice differentiable at* $x \in \Omega$ if

- f is differentiable in an open ball $B_\varepsilon(x)$ around x and
- the function $Df : B_\varepsilon(x) \to \mathcal{L}(X, V)$ is differentiable at x.

We say that f is twice differentiable on Ω if f is twice differentiable at every $x \in \Omega$.

The Second Derivative

2.5.1. Definition. Let X, V be finite-dimensional normed vector spaces and $\Omega \subset X$ an open set. A function $f : \Omega \to V$ is said to be *twice differentiable at* $x \in \Omega$ if

- f is differentiable in an open ball $B_\varepsilon(x)$ around x and
- the function $Df : B_\varepsilon(x) \to \mathcal{L}(X, V)$ is differentiable at x.

We say that f is twice differentiable on Ω if f is twice differentiable at every $x \in \Omega$.

Comment.

The derivative of Df (if it exists) is a map

$$D(Df) =: D^2 f : \Omega \to \mathcal{L}(X, \mathcal{L}(X, V)).$$

Hessian

The Hessian of f is

$$D(\nabla f)|_x = \begin{pmatrix} \left.\frac{\partial^2 f}{\partial x_1 \partial x_1}\right|_x & \left.\frac{\partial^2 f}{\partial x_2 \partial x_1}\right|_x & \cdots & \left.\frac{\partial^2 f}{\partial x_n \partial x_1}\right|_x \\ \vdots & \vdots & & \vdots \\ \left.\frac{\partial^2 f}{\partial x_1 \partial x_n}\right|_x & \left.\frac{\partial^2 f}{\partial x_2 \partial x_n}\right|_x & \cdots & \left.\frac{\partial^2 f}{\partial x_n \partial x_n}\right|_x \end{pmatrix} \in \mathrm{Mat}(n \times n; \mathbb{R})$$

Hessian

The Hessian of f is

$$D(\nabla f)|_x = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}\Big|_x & \frac{\partial^2 f}{\partial x_2 \partial x_1}\Big|_x & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1}\Big|_x \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}\Big|_x & \frac{\partial^2 f}{\partial x_2 \partial x_n}\Big|_x & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n}\Big|_x \end{pmatrix} \in \mathrm{Mat}(n \times n; \mathbb{R})$$

$$\begin{aligned} D^2 f|_x h &= D((\,\cdot\,)^T \circ \nabla f(x))h = D(\,\cdot\,)^T|_{\nabla f(x)} \circ D(\nabla f)|_x h \\ &= (\,\cdot\,)^T \circ D(\nabla f)|_x h = (\mathrm{Hess}\, f(x) h)^T. \end{aligned}$$

$$(D^2 f|_x h)\tilde{h} = (\mathrm{Hess}\, f(x) h)^T \tilde{h} = \langle \mathrm{Hess}\, f(x) h, \tilde{h} \rangle$$

$$D^2 f|_x \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \qquad (h, \tilde{h}) \mapsto \langle \mathrm{Hess}\, f(x) h, \tilde{h} \rangle$$

The Second Derivative as a Bilinear Map

Recall that

The derivative of Df (if it exists) is a map

$$D(Df) =: D^2 f : \Omega \to \mathcal{L}(X, \mathcal{L}(X, V)).$$

Let

$$L \in \mathcal{L}(X, \mathcal{L}(X, V))$$

Then

$$Lx_1 \in \mathcal{L}(X, V)$$
$$(Lx_1)(x_2) \in V.$$

The Second Derivative as a Bilinear Map

To $L \in \mathcal{L}(X, \mathcal{L}(X, V))$ we can associate a map $\widetilde{L}\colon X \times X \to V$

$$\widetilde{L}(x_1, x_2) := (Lx_1)(x_2)$$

We can check this map is bilinear

$$\begin{aligned}
\widetilde{L}(x_1, x_2 + x_2') &= (Lx_1)(x_2 + x_2') = (Lx_1)(x_2) + (Lx_1)(x_2') \\
&= \widetilde{L}(x_1, x_2) + \widetilde{L}(x_1, x_2'), \\
L(x_1, \lambda x_2) &= (Lx_1)(\lambda x_2) = \lambda (Lx_1)(x_2) = \lambda \widetilde{L}(x_1, x_2)
\end{aligned}$$

because $Lx_1 \in \mathcal{L}(X, V)$ is linear. Furthermore, since $L \in \mathcal{L}(X, \mathcal{L}(X, V))$,

$$\begin{aligned}
\widetilde{L}(x_1 + x_1', x_2) &= (L(x_1 + x_1'))(x_2) = (Lx_1 + Lx_1')(x_2) \\
&= (Lx_1)(x_2) + (Lx_1')(x_2) = \widetilde{L}(x_1, x_2) + \widetilde{L}(x_1', x_2), \\
\widetilde{L}(\lambda x_1, x_2) &= (\lambda Lx_1)(x_2) = \lambda (Lx_1)(x_2) = \lambda \widetilde{L}(x_1, x_2).
\end{aligned}$$

Multilinear Maps

The set of multilinear maps from X to V is denoted by

$$\mathcal{L}^{(n)}(X, V) := \Big\{ L\colon X \times \cdots \times X \to V \colon L \text{ linear in each component} \Big\}.$$

Example. Let $X = \mathbb{R}^n$ and $V = \mathbb{R}$. Then we have seen that

$$\mathcal{L}^{(2)}(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R}) \cong \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R})).$$

We know that $\mathcal{L}(\mathbb{R}^n, \mathbb{R}) = (\mathbb{R}^n)^* \cong \mathbb{R}^n$, so we have

$$\mathcal{L}^{(2)}(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R}) \cong \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n) \cong \operatorname{Mat}(n \times n, \mathbb{R}).$$

Multilinear Maps

Suppose

$$A \in \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}))$$

then

$$\mathrm{Mat}(n \times n, \mathbb{R}) \cong \mathcal{L}^{(2)}(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R}) \qquad \text{via} \qquad A \leftrightarrow \langle \cdot \, , A(\cdot) \rangle$$

Schwarz's Theorem

2.5.5. Schwarz's Theorem. Let X, V be normed vector spaces and $\Omega \subset X$ an open set. Let $f \in C^2(\Omega, V)$. Then $D^2 f|_x \in \mathcal{L}^{(2)}(X \times X, V)$ is symmetric for all $x \in \Omega$, i.e.,

$$D^2 f(u, v) = D^2 f(v, u), \qquad \text{for all } u, v \in X.$$

which means

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$$


Thank You!