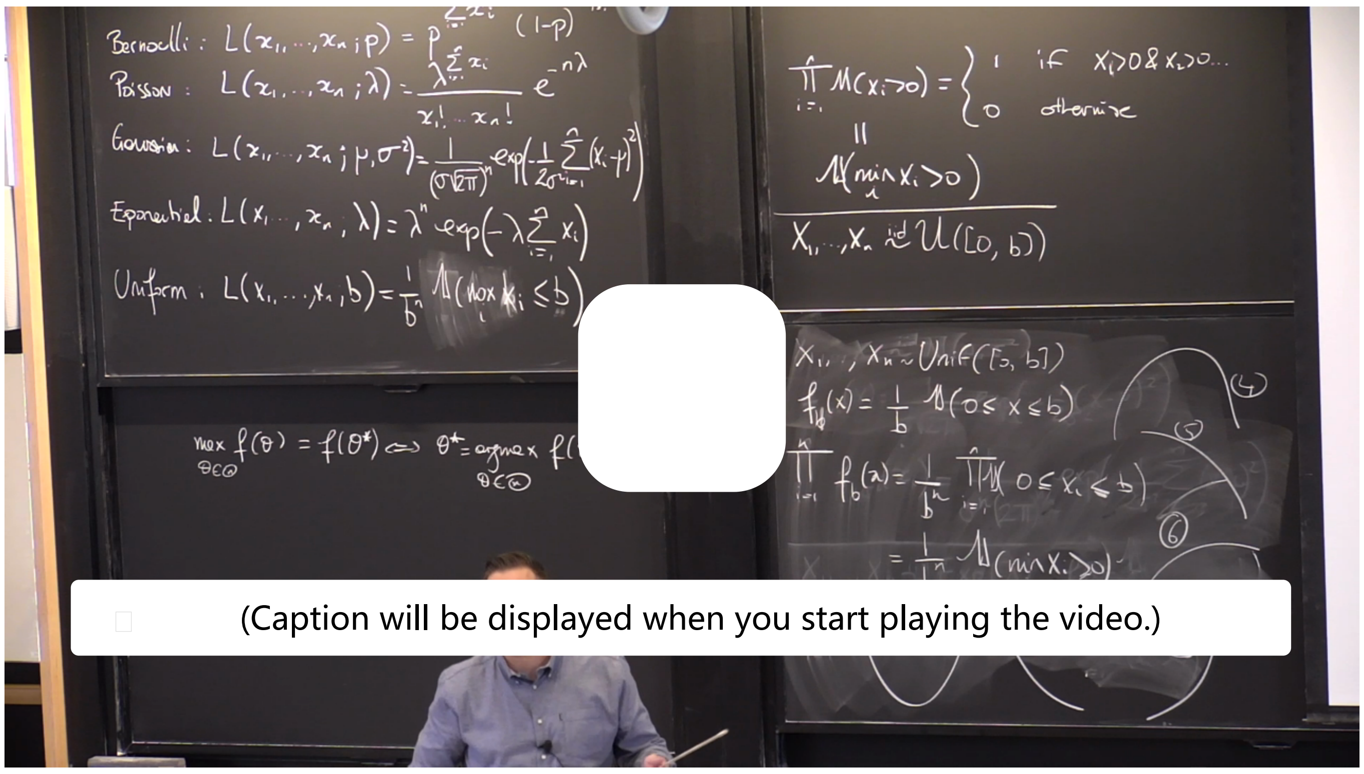
8. Review: Gradients and Hessians; Concavity in Higher dimensions

Concavity in Higher Dimensions: Gradients, Hessians, Semi-Definiteness



Start of transcript. Skip to the end.

So here, you can see we have one example where

the function we're looking at is actually

a function of two parameters-- one living on the real line

and one living on the positive real line.

And so we're going to have to talk about convex and concave

视频

下载视频文件

字幕

下载 SubRip (.srt) file

下载 Text (.txt) file

Multivariable Calculus Review: Compute the Gradient

1/1 point (graded)

Let

$$\begin{aligned} f: \quad & \mathbb{R}^d \quad \to \mathbb{R} \\ \theta = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_d \end{pmatrix} & \mapsto f(\theta) . \end{aligned}$$

denote a **differentiable** function. The **gradient** of f is the vector-valued function

$$\begin{aligned} \nabla f: \quad & \mathbb{R}^d \quad \to \mathbb{R}^d \\ \theta = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_d \end{pmatrix} & \mapsto \left. \begin{pmatrix} \frac{\partial f}{\partial \theta_1} \\ \frac{\partial f}{\partial \theta_2} \\ \vdots \\ \frac{\partial f}{\partial \theta_d} \end{pmatrix} \right|_{\theta} . \end{aligned}$$

Consider $f(\theta) = -c_1\theta_1^2 - c_2\theta_2^2 - c_3\theta_3^2$ where $c_1, c_2, c_3 > 0$ are positive real numbers.

Compute the gradient ∇f.

(Enter your answer as a vector, e.g., type **[3,2,x]** for the vector $\begin{pmatrix} 3 \\ 2 \\ x \end{pmatrix}$. Note the square brackets, and commas as separators. Enter **c_i** for c_i, **theta_i** for θ_i.)

$\nabla f =$ [-2*c_1*theta_1,-2*c_2*theta_2,-2*c_3*theta_3]

Answer: [-2*c_1*theta_1,-2*c_2*theta_2,-2*c_3*theta_3]

STANDARD NOTATION

Solution:

$$\begin{aligned} f(\theta) &= -c_1\theta_1^2 - c_2\theta_2^2 - c_3\theta_3^2 \\ \nabla f(\theta) &= \left.\begin{pmatrix} \frac{\partial f}{\partial \theta_1} \\ \frac{\partial f}{\partial \theta_2} \\ \frac{\partial f}{\partial \theta_3} \end{pmatrix}\right|_{\theta} = \begin{pmatrix} -2c_1\theta_1 \\ -2c_2\theta_2 \\ -2c_3\theta_3 \end{pmatrix}. \end{aligned}$$

提交 你已经尝试了1次（总共可以尝试3次）

Answers are displayed within the problem

Multivariable Calulus Review: Compute the Hessian Matrix

1/1 point (graded)
As above, let

$$f:\ \mathbb{R}^d \to \mathbb{R} \quad \theta = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_d \end{pmatrix} \mapsto f(\theta).$$

denote a **twice-differentiable** function.

The **Hessian** of f is the matrix

$$\mathbf{H} f:\ \mathbb{R}^d \to \mathbb{R}^{d\times d}$$

whose entry in the i-th row and j-th column is defined by

$$(\mathbf{H} f)_{ij} := \frac{\partial^2}{\partial\theta_i\partial\theta_j} f, \quad 1 \le i,j \le d.$$

Consider the same function $f(\theta) = -c_1\theta_1^2 - c_2\theta_2^2 - c_3\theta_3^2$ where $c_1, c_2, c_3 > 0$ as in the previous problem. Compute the Hessian matrix $\mathbf{H} f$.

(Enter your answer as a matrix, e.g. by typing **[[1,2],[5*x,y-1]]** for the matrix $\begin{pmatrix} 1 & 2 \\ 5x & y-1 \end{pmatrix}$. Note the square brackets, and commas as separaters.)

$\mathbf{H} f =$ [[-2*c_1,0,0],[0,-2*c_2,0],[0,0,-2*c_3]]

Answer: [[-2*c_1,0,0],[0,-2*c_2,0],[0,0,-2*c_3]]

STANDARD NOTATION

Solution:

Recall from the previous problem:

$$
\begin{aligned}
f(\theta) &= -c_1\theta_1^2 - c_2\theta_2^2 - c_3\theta_3^2 \\
\nabla f(\theta) &= \left.\begin{pmatrix} \frac{\partial f}{\partial \theta_1} \\ \frac{\partial f}{\partial \theta_2} \\ \frac{\partial f}{\partial \theta_3} \end{pmatrix}\right|_{\theta} = \begin{pmatrix} -2c_1\theta_1 \\ -2c_2\theta_2 \\ -2c_3\theta_3 \end{pmatrix}.
\end{aligned}
$$

One way to compute the Hessian is to start will in j-th column of the Hessian matrix by the gradient of the j-th component of ∇f. We obtain:

$$
\begin{aligned}
f(\theta) &= -c_1\theta_1^2 - c_2\theta_2^2 - c_3\theta_3^2 \\
\mathbf{H}f(\theta) &= \begin{pmatrix} | & | & | \\ \nabla(-2c_1\theta_1) & \nabla(-2c_2\theta_2) & \nabla(-2c_3\theta_3) \\ | & | & | \end{pmatrix} \\
&= \begin{pmatrix} -2c_1 & 0 & 0 & \\ 0 & 2c_2 & 0 & \\ 0 & 0 & & -2c_3 \end{pmatrix}.
\end{aligned}
$$

提交 你已经尝试了2次（总共可以尝试3次）

Answers are displayed within the problem

Semi-Definiteness

2/3 points (graded)

A symmetric (real-valued) $d \times d$ matrix $\mathbf{A}$ is **positive semi-definite** if

$$\mathbf{x}^T \mathbf{A}\,\mathbf{x} \geq 0 \qquad \text{for all } \mathbf{x} \in \mathbb{R}^d.$$

If the inequality above is strict, i.e. if $\mathbf{x}^T \mathbf{A}\,\mathbf{x} > 0$ for all non-zero vectors $\mathbf{x} \in \mathbb{R}^d$, then $\mathbf{A}$ is **positive definite** .

Analogously, a symmetric (real-valued) $d \times d$ matrix $\mathbf{A}$ is **negative semi-definite** (*resp.* **negative definite**) if $\mathbf{x}^T \mathbf{A}\,\mathbf{x}$ is **non-positive** (*resp.* **negative**) for all $\mathbf{x} \in \mathbb{R}^d - \{\mathbf{0}\}$.

Note that by definition, positive (or negative) definiteness implies positive (or negative) semi-definiteness.

Consider the same function as in the problems above:

$$f(\theta) = -c_1\theta_1^2 - c_2\theta_2^2 - c_3\theta_3^2 \quad \text{where } c_1, c_2, c_3 > 0.$$

Compute $\mathbf{x}^T (\mathbf{H}f)\,\mathbf{x}$ where $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$.

$\mathbf{x}^T (\mathbf{H}f)\,\mathbf{x} =$ -2*c_1*x_1^2 - 2*c_2*x_2^2 - 2*c_3*x_3^2

Answer: -2*c_1*x_1^2-2*c_2*x_2^2-2*c_3*x_3^2

$-2 \cdot c_1 \cdot x_1^2 - 2 \cdot c_2 \cdot x_2^2 - 2 \cdot c_3 \cdot x_3^2$

The matrix $\mathbf{H}f$ is (Choose all that apply.)

- [ ] positive semi-definite
- [ ] positive definite
- [ ] negative semi-definite
- [x] negative definite

Hence, the function f is
(Choose all that apply.)

- [x] concave
- [x] strictly concave
- [ ] convex
- [ ] strictly convex

Solution:

Recall from the previous problem that

$$\mathbf{H}f(\theta) = \begin{pmatrix} -2c_1 & 0 & 0 \\ 0 & 2c_2 & 0 \\ 0 & 0 & -2c_3 \end{pmatrix}.$$

Then

$$\begin{aligned} \mathbf{x}^T(\mathbf{H}f)\,\mathbf{x} &= \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} -2c_1 & 0 & 0 & \\ 0 & 2c_2 & 0 & \\ 0 & 0 & & 2c_3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\ &= -2c_1x_1^2 - 2c_2x_2^2 - 2c_3x_3^2 < 0. \end{aligned}$$

Since $c_1, c_2, c_3 > 0$, this means the $\mathbf{H}f$ is negative definite, (also negative semi-definite), and hence f is strictly concave (also concave).

提交 你已经尝试了2次（总共可以尝试2次）

Answers are displayed within the problem

讨论

显示讨论

主题: Unit 3 Methods of Estimation:Lecture 9: Introduction to Maximum Likelihood Estimation / 8. Review: Gradients and Hessians; Concavity in Higher dimensions