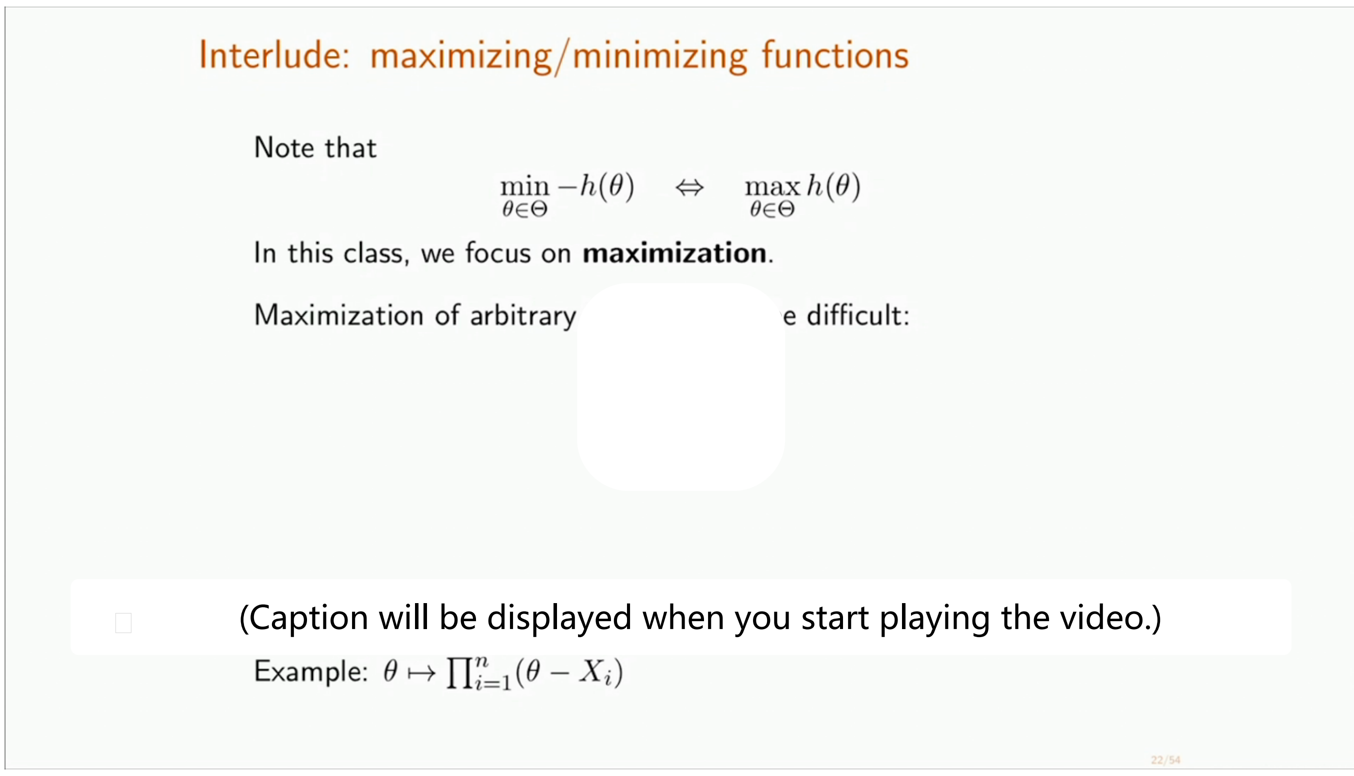
# 6. Interlude: Minimizing and Maximizing Functions

## Concavity in 1 dimension




Start of transcript. Skip to the end.

So the first thing is, in this class,

we talk a lot about maximization because we

have the maximum likelihood estimator.

Now, maximizing or minimizing a function

is actually referred often to as optimization.

It's an actual entire field.

And actually, it's been really revived by machine learning,

and if you go to NIPS, which is the huge

**视频**

下载视频文件

**字幕**

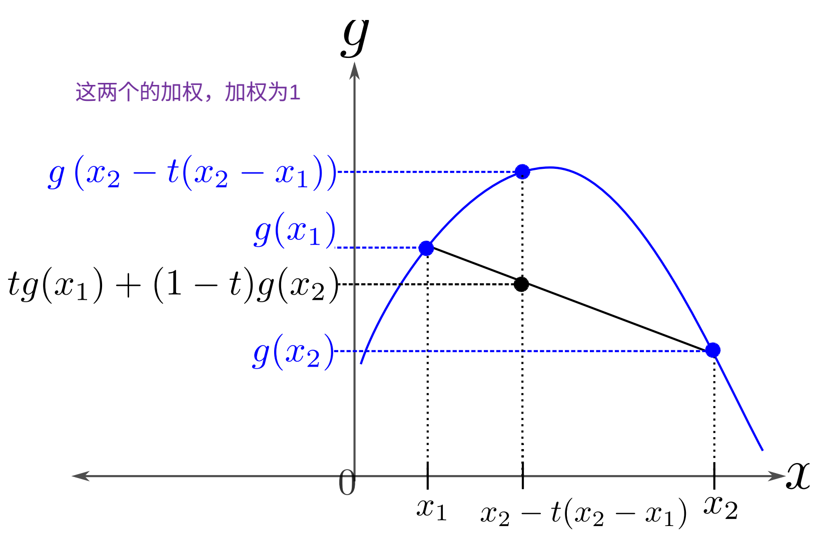
下载 SubRip (.srt) file

下载 Text (.txt) file

A function $g: I \to \mathbb{R}$ is **concave** (or concave down), where $I$ is an interval, if for all pairs of real numbers $x_1 < x_2 \in I$

$$g\left(tx_1 + (1-t)\,x_2\right) \geq t g\left(x_1\right) + (1-t)\,g\left(x_2\right) \qquad \text{for all } 0 < t < 1.$$

Geometrically, this means that for $x_1 < x < x_2$, the graph of $g$ is **above** the secant line connecting the two points $\left(x_1, g\left(x_1\right)\right)$ and $\left(x_2, g\left(x_2\right)\right)$.




At $x = x_2 - t\left(x_2 - x_1\right) = tx_1 + (1-t)\,x_2$, the $y$-value of the graph of $g$ is $g\left(x\right) = g\left(tx_1 + (1-t)\,x_2\right)$, while the $y$-value of the secant line is $tg\left(x_1\right) + (1-t)\,g\left(x_2\right)$.

If the inequality is strict, i.e. if

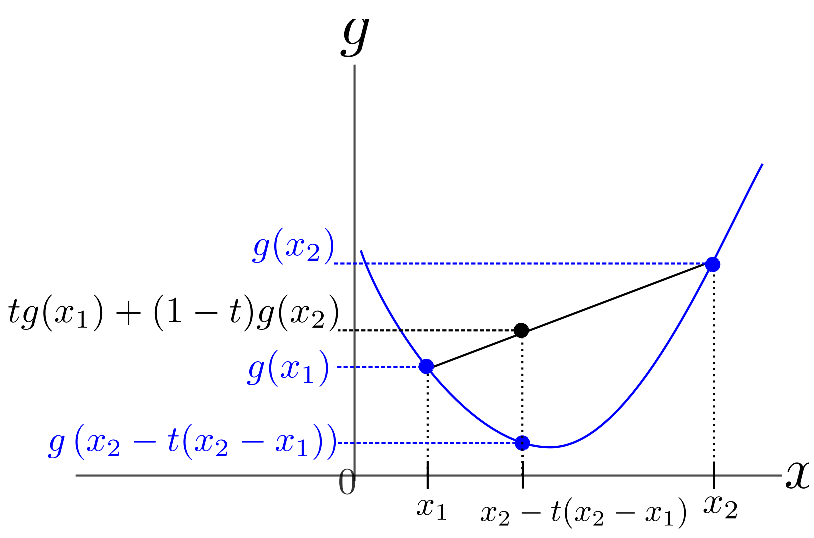
$$g\left(tx_1 + (1-t)\,x_2\right) > t g\left(x_1\right) + (1-t)\,g\left(x_2\right) \qquad \text{for all } 0 < t < 1.$$

then $g$ is **strictly concave** .

The definition for **(strictly) convex** is analogous. A function $g : I \to \mathbb{R}$ is **convex** (or concave up), where $I$ is an interval, if for all pairs of real numbers $x_1 < x_2 \in I$

$$g(tx_1 + (1-t)x_2) \leq tg(x_1) + (1-t)g(x_2) \qquad \text{for all } 0 < t < 1.$$

Geometrically, this means that for $x_1 < x < x_2$, the graph of $g$ is **below** the secant line connecting the two points $(x_1, g(x_1))$ and $(x_2, g(x_2))$.




At $x = x_2 - t(x_2 - x_1) = tx_1 + (1-t)x_2$, the $y$-value of the graph of $g$ is $g(x) = g(tx_1 + (1-t)x_2)$, while the $y$-value of the secant line is $tg(x_1) + (1-t)g(x_2)$.

If the inequality is strict, i.e. if

$$g(tx_1 + (1-t)x_2) < tg(x_1) + (1-t)g(x_2) \qquad \text{for all } 0 < t < 1.$$

then $g$ is **strictly convex** .

If in addition $g$ is twice differentiable in the interval $I$, i.e. $g''(x)$ exists for all $x \in I$, then $g$ is

- **concave** if and only if $g''(x) \leq 0$ for all $x \in I$;
- **strictly concave** if $g''(x) < 0$ for all $x \in I$;
- **convex** if and only if $g''(x) \geq 0$ for all $x \in I$;
- **strictly convex** if $g''(x) > 0$ for all $x \in I$;

## 讨论

显示讨论

**主题:** Unit 3 Methods of Estimation:Lecture 9: Introduction to Maximum Likelihood Estimation / 6. Interlude: Minimizing and Maximizing Functions