PRACTICAL EXERCISES
THURSDAY 18/4

https://finans.dk/erhverv/ECE17026456/branden-paa-boersen-betoed-oeget-eksponering-til-airtoxreklame/

CLASS 9

OPTIMISATION AND MAXIMUM LIKELIHOOD ESTIMATION

STUDENT INSTRUCTOR | PERNILLE HØJLUND BRAMS (AU650502@CAS.AU.DK)
COURSE | METHODS 2: THE GENERAL LINEAR MODEL

AGENDA

- Recap +:
  - What is optimisation?
  - What is the Maximum Likelihood Estimate?
- Break
- Exercises 📝 + portfolio 2 work
- More recap if time

https://www.scribbr.com/statistics/normal-distribution/

https://da.wikipedia.org/wiki/B%C3%B8rsen

OPTIMISATION 1

TAKE A FEW MINUTES IN GROUPS:

1) What is f’(x) and what can we use it for?
2) How can you draw a case of f’(x) = 0 and what can we use f’(x) = 0 for?

OPTIMISATION 1

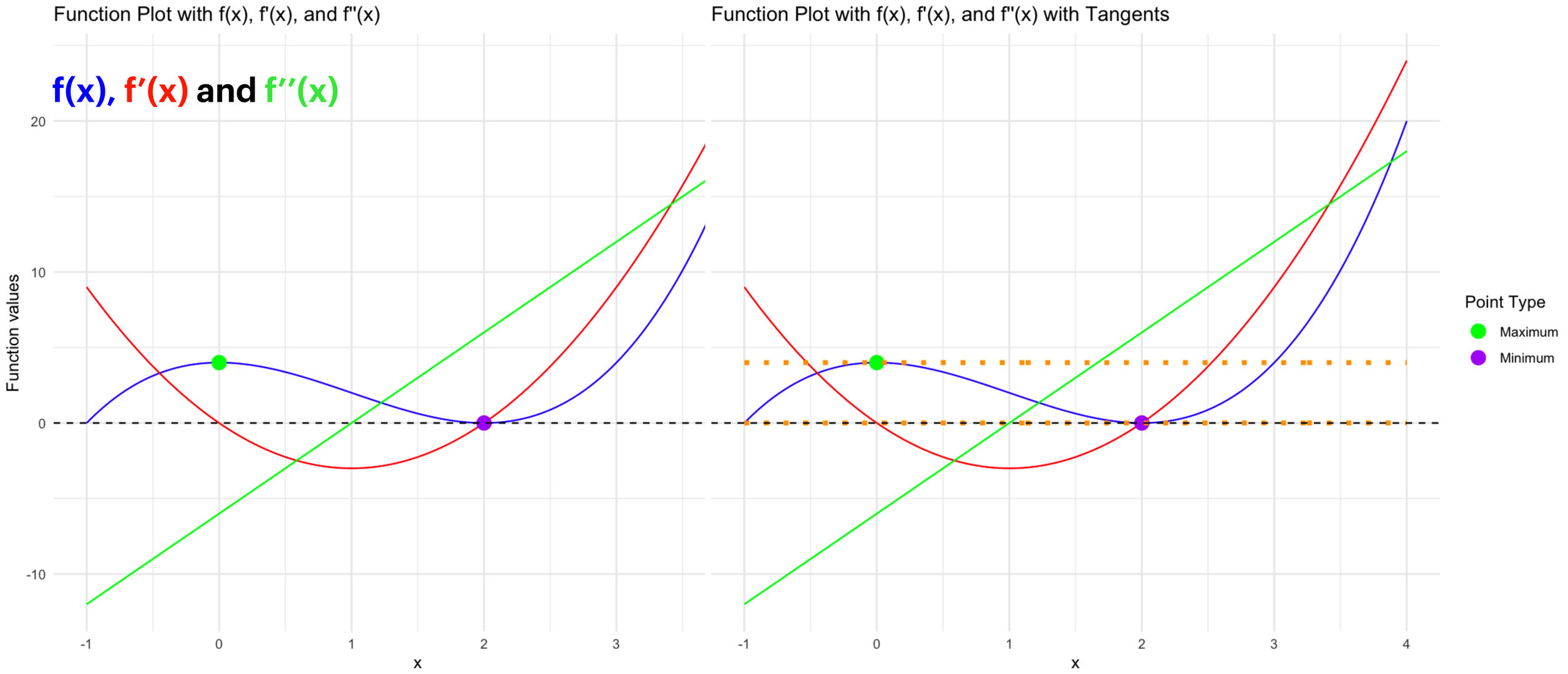
TAKE A FEW MINUTES IN GROUPS:

1) **What is f’(x) and what can we use it for?**
2) **How can you draw a case of f’(x) = 0 and what can we use f’(x) = 0 for?**

f'(x) is the first derivative of a function f(x), used to determine the rate of change of f and identify critical points where the slope is zero. To illustrate f'(x) = 0, we can plot f(x) and mark points where its tangent is horizontal; these points help identify potential maxima, minima, or inflection points in f(x).

LET'S VISUALISE IT: WITHOUT TANGENTS

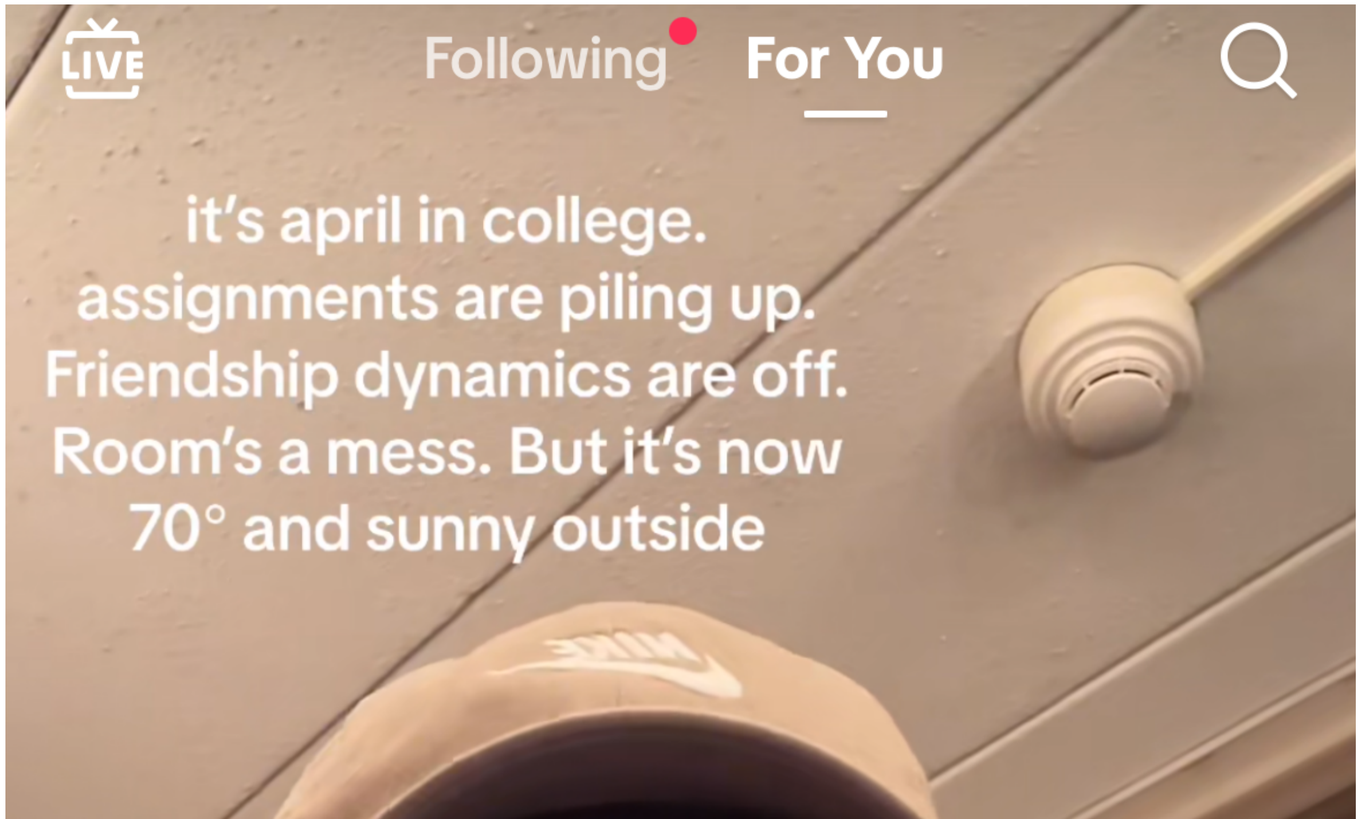
WITH TANGENTS (IN ORANGE)



OPTIMISATION 2

1 MINUTE ALONE —> 2 MINUTES IN GROUPS:

IN YOUR OWN WORDS: WHAT DOES OPTIMISATION MEAN?



OPTIMISATION 2

1 MINUTE ALONE —> 2 MINUTES IN GROUPS:

IN YOUR OWN WORDS: WHAT DOES OPTIMISATION MEAN?

OPTIMISATION, IN SIMPLE TERMS, IS THE PROCESS OF MAKING SOMETHING AS EFFECTIVE OR FUNCTIONAL AS POSSIBLE. IN VARIOUS CONTEXTS, THIS USUALLY INVOLVES FINDING THE BEST SOLUTION FROM A SET OF AVAILABLE ALTERNATIVES. FOR EXAMPLE, IN MATHEMATICS OR COMPUTING, OPTIMISATION MIGHT INVOLVE FINDING THE MAXIMUM OR MINIMUM VALUE OF A FUNCTION, WHICH COULD REPRESENT COST, EFFICIENCY, SPEED, PROFIT, OR OTHER MEASURABLE FACTORS.

OPTIMISATION

Consider a function $f(\mathbf{x})$ of a vector $\mathbf{x}$. Optimization problems are concerned with the task of finding $\mathbf{x}^\star$ such that $f(\mathbf{x}^\star)$ is a local maximum (or minimum). In the case of maximization,

$$\mathbf{x}^\star = \text{argmax } f(\mathbf{x})$$

and in the case of minimization,

$$\mathbf{x}^\star = \text{argmin } f(\mathbf{x})$$

Most statistical estimation problems are optimization problems. For example, if f is the likelihood function and $\mathbf{x}$ is a vector of parameter values, then $\mathbf{x}^\star$ is the **maximum likelihood estimator (MLE)**, which has many nice theoretical properties.

OPTIMISATION

Consider a function $f(\mathbf{x})$ of a vector $\mathbf{x}$. Optimization problems are concerned with the task of finding $\mathbf{x}^\star$ such that $f(\mathbf{x}^\star)$ is a local maximum (or minimum). In the case of maximization,

x* = "x-star" is the "best" input value, giving us the highest/lowest output value

$$\mathbf{x}^\star = \operatorname{argmax} f(\mathbf{x})$$

and in the case of minimization,

$$\mathbf{x}^\star = \operatorname{argmin} f(\mathbf{x})$$

Most statistical estimation problems are optimization problems. For example, if f is the likelihood function and $\mathbf{x}$ is a vector of parameter values, then $\mathbf{x}^\star$ is the **maximum likelihood estimator (MLE)**, which has many nice theoretical properties.

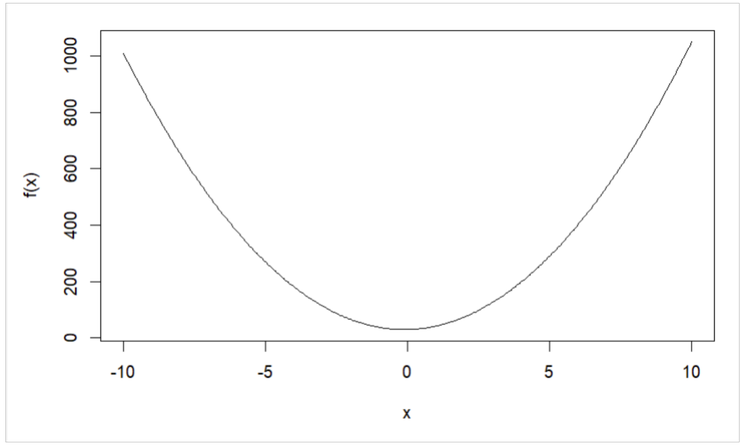
MLE seeks the x* that gives highest value of the likelihood function.
The values of x* that achieve this are known as 'maximum likelihood estimators'

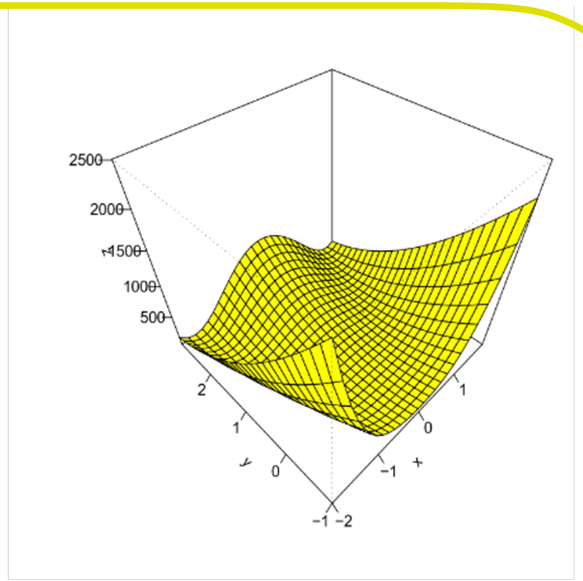
✨ MLE SUMMED UP ✨

MLE works by calculating the probability of occurrence for each data point (we call this the likelihood) for a model with a given set of parameters.

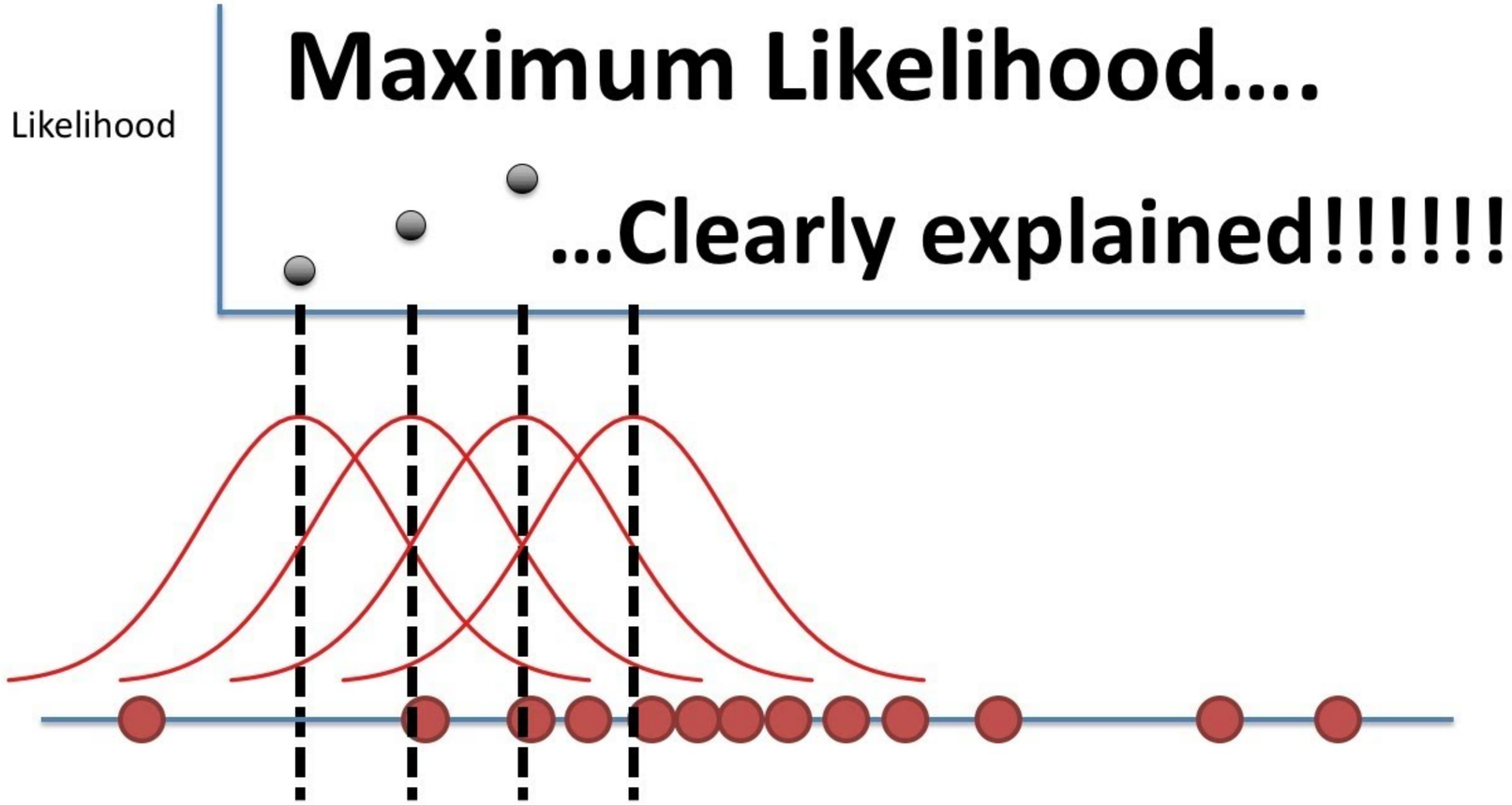
MLE is tied to optimisation because it fundamentally involves an optimisation problem.

OPTIMISATION





MLE CLEARLY EXPLAINED



THE SECOND DERIVATIVE RULE

Local maximum if f’(x) = 0 & f’’(x) = negative (concave down)

Local minimum if f’(x) = 0 & f’’(x) = positive (concave up)

Doing this example on the board using the second derivative test.
Steps:

Eq: $f(x) = x^2 + 3x - 8$

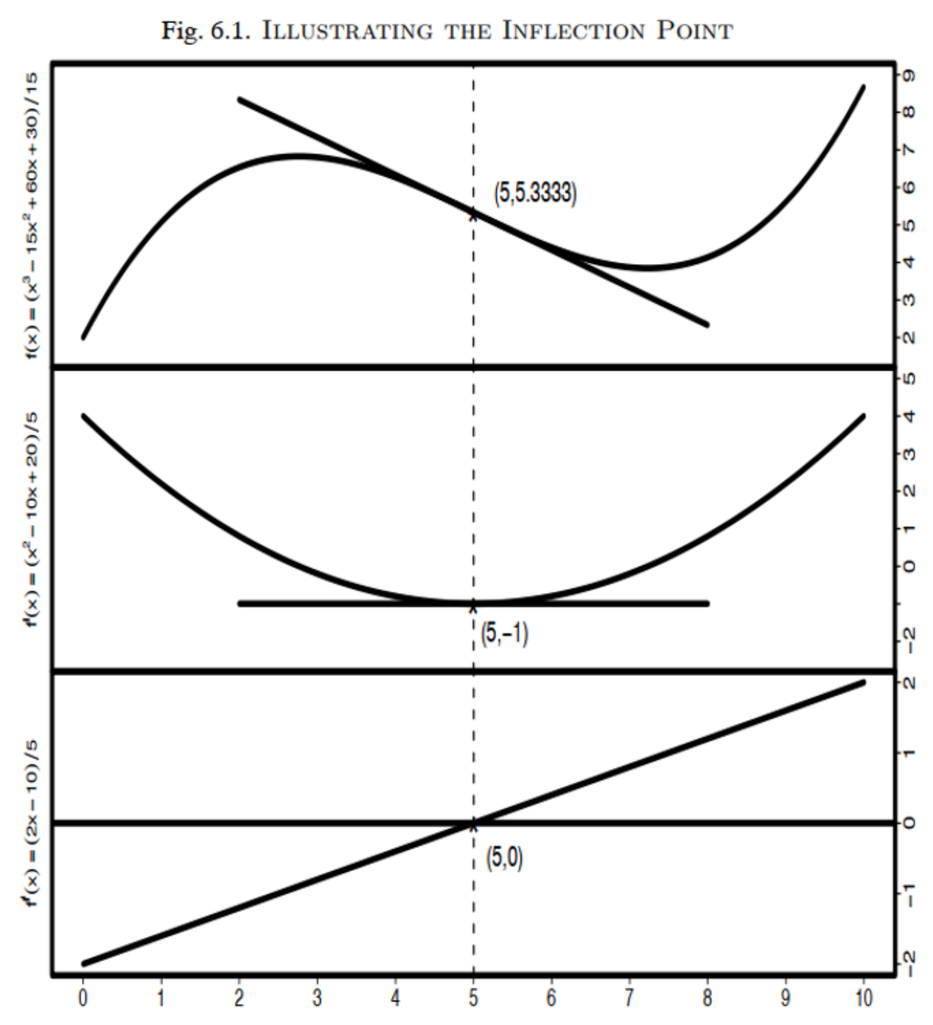
1. $f'(x) = ?$

2. When is $f'(x) = 0$? Call x for f’(x) = 0 the letter c.

3. $f''(x) = ?$

4. $f''(c) = ?$

5. “Since f’’(c) >/< 0 by the second derivative test, f(x) has a relative minimum at x = c”



THE SECOND DERIVATIVE RULE W. SOLUTION

Local maximum if f’(x) = 0 & f’’(x) = negative (concave down)

Local minimum if f’(x) = 0 & f’’(x) = positive (concave up)

Doing this example on the board using the second derivative test.
Steps: w. solution

Eq: $f(x) = x^2 + 3x - 8$

1. $f'(x) = ?$ $f'(x) = 2x + 3$

2. When is $f'(x) = 0$? Call x for f’(x) = 0 the letter c. c = -3/2

3. $f''(x) = ?$ $f''(x) = 2$

4. $f''(c) = ?$ $f''(c) = f''(-3/2) = 2$

5. “Since f’’(c) > 0 by the second derivative test, f(x) has a local minimum at x = -3/2”

Fig. 6.1. Illustrating the Inflection Point

WHY 2ND DERIVATIVE, OPTIMISATION AND MLE IN THE SAME CLASS???💀

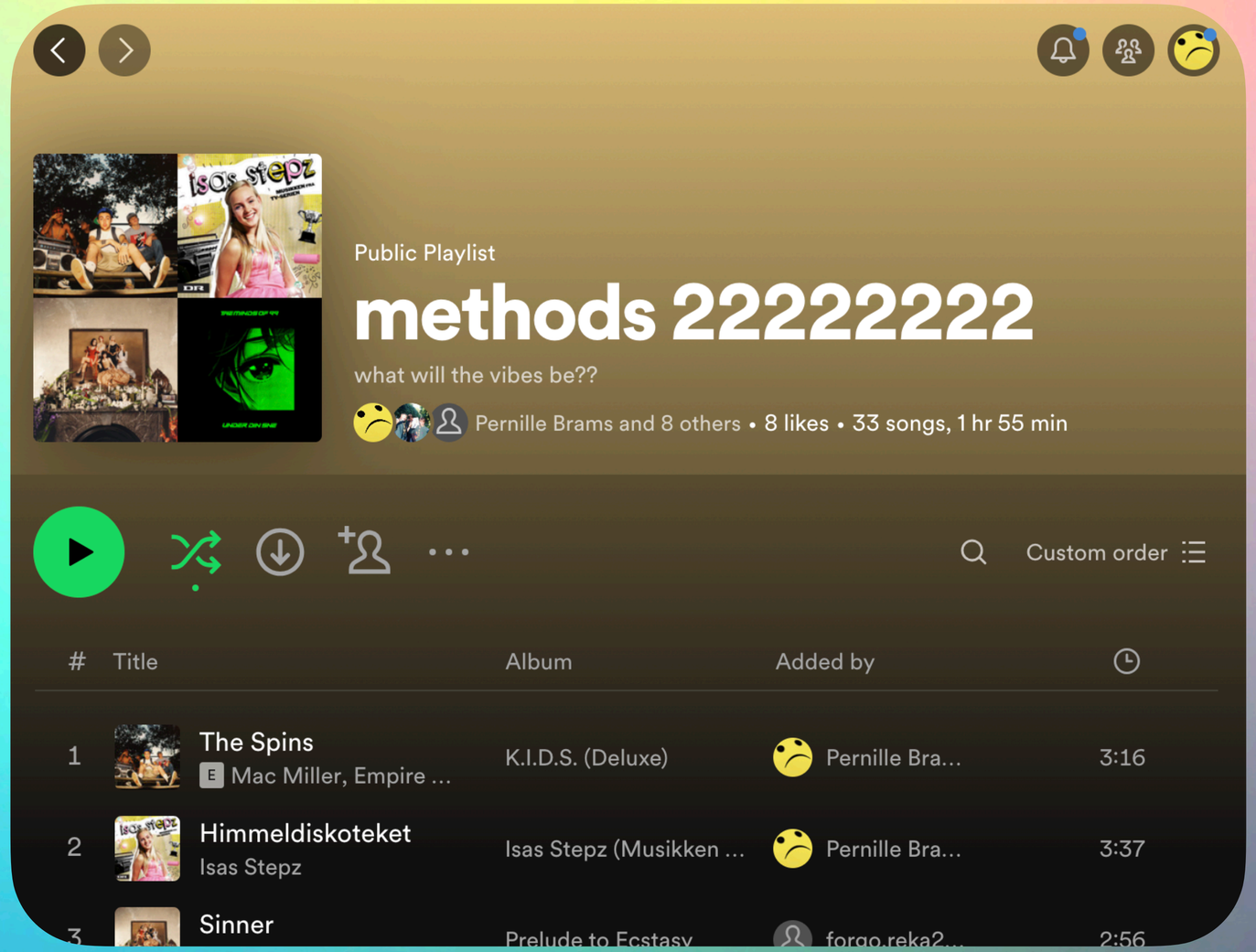
1) Understanding the second derivative test provides math tools to engage optimisation problems (finding extrema)

2) MLE is a practical application of optimisation. We try to find parameters that maximise the likelihood function (making the observed data most probable under the model)

3) Goal is to learn how to perform optimisation, but also to understand why certain solutions are preferred (e.g. via MLE) and how to verify the solutions (e.g. using 2nd derivative test)

EXERCISES

See .Rmd markdown for this class (9) on Git, and if extra time: exercises for review (**deadline for submission if you want feedback is tomorrow, 19/4**)



Collab: https://open.spotify.com/playlist/5UUiKD15vyFwymQ4qLur9V?si=cf80f6c72721427c&pt=a27065eeba53e23fe2dd160612aa5598