Today...

1. Survey
2. Fitting a line to data
3. Chain Rule in Applications
I think that the lecture pace is

A. Way too fast
B. Too fast
C. Just right
D. Too slow
E. Way too slow
I think that the lecture clarity is
A. Way too clear
B. Too clear
C. Just right
D. Too confusing
E. Way too confusing
I think that the examples in lecture are
A. Way too many
B. Too many
C. Just right
D. Too few
E. Way too few
Survey

I think that the life science applications are
A. Way too many
B. Too many
C. Just right
D. Too few
E. Way too few
Survey

I think that Cole is
A. Way too friendly, approachable, helpful
B. Above average
C. Just right as my Math 102 instructor
D. Below average
E. Way too unfriendly, not approachable, and useless
Survey

I think that the notes I take in lecture are
A. Way too useful
B. Too useful
C. Just fine
D. Too useless
E. Way too useless
I would prefer if Cole used
A. The blackboard
B. The document camera
C. Slides
D. Combination of B and C and gave up on blackboard drawings
Survey

Do I take note of, and review the material that I don’t understand?

A. Yes, but I still don’t get it.
B. Yes, and I figure it out,
C. Yes, and I sometimes figure it out.
D. No
Survey

I don’t like to ask questions in class because
A. I don’t want Cole to know that I don’t understand
B. I don’t want my classmates or my friends to know that I don’t understand
C. I don’t have time to formulate questions before the class moves on
D. A or B or C or all three
E. I understand and I don’t have any questions to ask
Survey

Compared to my other classes, I think that Math 102 is
   A. Way too hard
   B. Too hard
   C. Just right
   D. Too easy
   E. Way too easy
Survey

Compared to my other classes, I think that Math 102 requires
A. Way too much work
B. Too much work
C. Just the right amount of work
D. Too little work
E. Way too little work
We’re in this together...

- Office Hours: https://wiki.math.ubc.ca/mathbook/M102/Section_links/Section_106/2016#Links_to_Lecture_Slides
- Email: zmurchok@math.ubc.ca
Fitting a line to data

A statistical model is a function that approximates a set of data.

Example: linear model $y = ax$ attempts to explain how $\Delta$ length $= a \times$ concentration.
Fitting a line to data

[Graph showing a scatter plot with concentration (µM) on the x-axis and Δ length on the y-axis.]

[Link to relevant online resource provided]

https://wiki.math.ubc.ca/mathbook/M102/Course_notes/Fitting_data_-_least_squares
Fitting a line without intercept to data

Q1. Suppose you have data \((x_1, y_1), (x_2, y_2), (x_3, y_3)\), and wish to fit a line \(y = ax\) through these points.

A residual for \((x_i, y_i)\) is

A. \(r_i = y_i^2 + x_i^2\)

B. \(r_i = y_i - ax_i\)

C. \(r_i = a^2(y_i^2 + x_i^2)\)

D. \(r_i = y_i - x_i\)

E. \(r_i = x_i - y_i\)
Fitting a line without intercept to data

Q2. Suppose you have data \((x_1, y_1), (x_2, y_2), (x_3, y_3)\), and wish to fit a line \(y = ax\) through these points.

Which graph below shows the residuals, \((y_i - ax_i)\)?

Graph B
Fitting a line without intercept to data

Q3. To get the best fit line $y = ax$ we should minimize

A. $f(a) = \sum_{i=1}^{3} (x_i - y_i)^2$

B. $f(a) = \sum_{i=1}^{3} (y_i - ax_i)^2$

C. $f(a) = \sum_{i=1}^{3} |y_i - ax_i|$

D. $f(a) = \sum_{i=1}^{3} (y_i - ax_i)$

Minimize the sum of square residuals. We minimize the square of the residual to
Fitting a line without intercept to data

Q4. The derivative of

\[(y_i - ax_i)^2\]

is

A. \(-2a(y_i - ax_i)\)

B. \(-2x_i(y_i - ax_i) = -2(x_iy_i - ax_i^2)\) (Chain rule)

C. 0

D. \(-2ax_i\)
Fitting a line without intercept to data

Q5. The derivative of \( f(a) = \sum_{i=1}^{3} (y_i - ax_i)^2 \) is

A. \( f'(a) = -2 \left( \sum_{i=1}^{3} y_i x_i - a \left( \sum_{i=1}^{3} x_i \right)^2 \right) \)

B. \( f'(a) = 2 \left( \sum_{i=1}^{3} y_i x_i - a \sum_{i=1}^{3} x_i^2 \right) \)

C. \( f'(a) = -2 \left( \sum_{i=1}^{3} y_i x_i - a \sum_{i=1}^{3} x_i^2 \right) \)
Fitting a line without intercept to data

\[ f'(a) = -2 \left( \sum_{i=1}^{3} y_i x_i - a \sum_{i=1}^{3} x_i^2 \right) = 0 \]

for

\[ a = \frac{\sum_{i=1}^{3} y_i x_i}{\sum_{i=1}^{3} x_i^2} \]
Summary

To find the line, \( y = ax \), of best fit to with \( n \) data points \((x_i, y_i)\), minimize the sum of square residuals:

\[
f(a) = \sum_{i=1}^{3} (y_i - ax_i)^2.
\]

Note that

\[
f'(a) = -2 \left( \sum_{i=1}^{n} y_ix_i - a \sum_{i=1}^{n} x_i^2 \right) = 0
\]

for

\[
a = \frac{\sum_{i=1}^{n} y_ix_i}{\sum_{i=1}^{n} x_i^2}
\]
Practice Problem

Find the line $y = ax$ of best fit to the data points in the table from WeBWorK 7 Question 5 (data.xlsx)

Hint: Use a spreadsheet and the formula

$$a = \frac{\sum_{i=1}^{n} y_i x_i}{\sum_{i=1}^{n} x_i^2}$$
Reminder: Chain Rule

If \( y = f(u) \) and \( u = g(x) \) are both differentiable functions and \( y = f(g(x)) \) is the composite function, then the \textit{chain rule} of differentiation states that

\[
y' = f'(u)g'(x) = f'(g(x))g'(x)
\]

or written another way

\[
\frac{dy}{dx} = \frac{dy}{du} \frac{du}{dx}
\]
Coffee Budget = Chain Rule \[ \frac{dy}{dx} = \frac{dy}{du} \frac{du}{dx} \]

rate of increase of money spent on coffee

= 

d rate of increase of price per cup 

\times

d rate of increase of cups I drink
A bear forages for salmon and berries. It spends fraction $x$ of its time foraging for food type $A$ (so $0 < x < 1$). Fraction $(1 - x)$ of time foraging for food type $B$, which is $N$ times more nutritious than $A$. Probability of finding food as a function of time spent

- for $A$: $x^\alpha$
- for $B$: $(1 - x)^\beta$

Maximize the expected nutritional value

$$V(x) = x^\alpha + N(1 - x)^\beta$$

for $\alpha = 0.4$, $\beta = 0.5$, and $N = 1$. 
Chain Rule

\[ \frac{dy}{dx} = \frac{dy}{du} \frac{du}{dx} \]

- Chapter 8 of the Course Notes is a good place to look for optimization problems involving the chain rule.
- These problems can be quite tricky! For those who don’t know, office hours are held in small classrooms—you can come work on problems/homework and ask questions!
Answers

1. B
2. C
3. B
4. B
5. C