# 8. Summary

Bookmark this page

## Summarize

**Big Picture**

Given a set of data points, $(x_i, y_i)$ the simplest possible relationship we can hope to find between the data is a linear one. That is, we want to find a line $y = ax + b$ that best fits these data. Our "variables" are now the $a$ and $b$, the slope and $y$-intercept that determine the linear relationship.

In order to find the best possible $a$ and $b$, we minimize an error function. The error function we defined squares the difference between the $y_i$ to the predicted value $ax_i + b$:

$$D(a,b) = \sum_{i=1}^{n} \left(y_i - (ax_i + b)\right)^2.$$

Thus finding the best fit requires minimizing this function of two variables $a$ and $b$, which is an (unconstrained) optimization problem!

**Mechanics**

To fit a collection of data $(x_i, y_i)$ to a straight line that minimizes the squares of the differences between the predicted $y$ values and the actual $y_i$, solve the system:

$$\left(\sum_{i=1}^{n} x_i^2\right) a + \left(\sum_{i=1}^{n} x_i\right) b = \sum_{i=1}^{n} x_i y_i \tag{4.218}$$

$$\left(\sum_{i=1}^{n} x_i\right) a + nb = \sum_{i=1}^{n} y_i. \tag{4.219}$$

where $n$ is the number of data points.

- To fit your data $(x_i, y_i)$ to a power law $y = cx^p$, first transform your data $(X_i, Y_i) = (\ln x_i, \ln y_i)$, then use least squares approximation. The power $p$ will be the slope.
- To fit your data $(x_i, y_i)$ to an exponential law $y = ce^{kx}$, first transform your data $(X_i, Y_i) = (x_i, \ln y_i)$. Then use least squares approximation. The exponential multiple $k$ is the slope of the least squares fitting to the transformed data.

**Ask Yourself**

**Why is it called least-squares approximation?**

In order to define "best fit" we have to make a decision about how to quantify the error. A common choice is to minimize the square of the difference between predicted and observed values. Since this method involves minimizing squares, the method is called "least-squares".

Hide

**What is the point of drawing a straight line through data?**

Fitting a straight line to data isn't just for making pretty pictures! Once we have a line that fits our data well, we can "extrapolate", which means we can make a prediction for what would have happened for values we weren't able to test. For example, if you find the line of best-fit between $x =$ heights of marathon runners and $y =$ marathon time, then you can use the line to predict the time for a new marathon runner, even if all you know is their height. Of course the prediction won't be perfect, but it's a good start. If you study statistics, you can learn how to create a "confidence interval" around the predicted

Calculator Hide Notes

In this sense, least squares approximation is an example of *machine learning*, the process by which a computer can be "trained" on some data (here it would be the $(x_i, y_i)$ pairs that we measure) and the computer "learns" to predict what would happen for other data (new values of $x_i$ outside of the training set).

Hide

## 8. Summary

Hide Discussion

**Topic:** Unit 3: Optimization / 8. Summary

**Add a Post**

Show all posts by recent activity

Something fascinating: transforming proxy into equality!
Community TA
5

Typos - [STAFF]
2

Previous

Next Up: Recitation 12: Fitting a parabola to data
11 min + 4 activities



# edX

About
Affiliates
edX for Business
Open edX
Careers
News

# Legal

Terms of Service & Honor Code
Privacy Policy
Accessibility Policy
Trademark Policy
Sitemap

# Connect

Calculator Hide Notes

Blog
Contact Us
Help Center
Media Kit
Donate



Calculator Hide Notes