Bradshaw's Bakery sold one customer 48 chocolate cookies and 108 oatmeal cookies for $110. The bakery also sold another customer 12 oatmeal cookies for $10. How much do the cookies cost?

This problem can be expressed as simultaneous equations:

We prefer to model linear equations using matrices:

This is a convenient because we can use arrays to represent matrices.

In ordinary arithmetic (without matrices), we can solve like this:

For example:

But there is no valid way to divide a matrix by another matrix!

Luckily, there is another way!

Using a special matrix – the inverse, we can achieve the same effect:

Here, the matrix is the identity matrix. It is the matrix equivalent to the number 1.

Of course, to findin using matrices, we must find the inverse matrix using techniques that we won’t go into today. The problem with this, is that *sometimes there is no inverse*. This happens when there are less equations than unknowns (underdetermined system), or more equations than unknowns (overdetermined system).

In these cases, where there is no unique solution, and no inverse, we must find the pseudoinverse, such that:

And so:

In this case,provides the ‘best-fit’ weightings such that:

An interesting property, which tonight’s worksheet points out, is that when we find the best-fit weightings, is at a minimum:

ONTO THE WORKSHEET ☺

Imagine we have an nth degree Polynomial, given by:

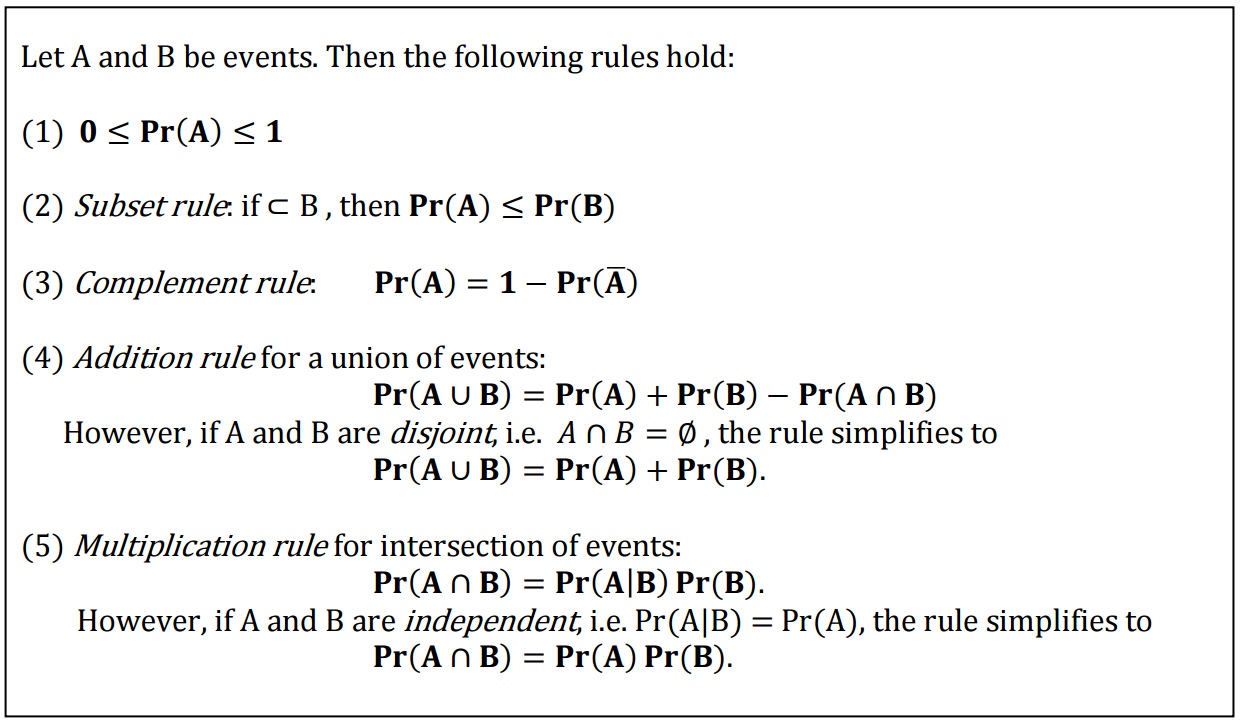
Where:

We want to find:

Such that the following is a minimum:

*The sum of the differences between estimated points and actual points.*

**Exercise 4**



Recall Bayes’ rule:

Show that the following is also true:

Recall the multiplication rule for intersections of events:

We now have an expression for! But it’s not quite there…

Consider the multiplication rule again:

If we change the event into:

Substituting the first result for into:

This is very close! This is what we want:

Consider the multiplication rule again:

And we’re done!

**Exercise 5:**

Suppose we want to fit the linear function, where is a scalar to a dataset.

Compute the derivative of the error with respect to:

Find the derivative using the chain rule:

Solve for: