Chapter 2 (continued):

Linear Approximation and Error Analysis

Prerequisites: Small increments from calculus of one variable. Triangle inequality. Maclaurin and Taylor series. Variance, standard deviation.

Small increments for one variable (review)

For now, we're going back to a function of one variable.

Suppose y depends on x through the function $y = f(x)$.

A small change Δx in x causes a small change Δy in y.

Here Δx and Δy are called *small increments.*

We want to develop an approximate way of calculating Δy that allows us to apply it easily in lots of situations; we could of course do this exactly by writing $\Delta y = f(x + \Delta x) - f(x)$, but we want to take a different approach here that uses the derivative.

Recall that $\frac{\Delta y}{\Delta x}$ becomes the derivative $\frac{dy}{dx}$ as $\Delta x \to 0$.

This means we can *approximate* $\frac{\Delta y}{\Delta x} \approx \frac{dy}{dx}$.

Multiplying this across by Δx gives $\Delta y \approx \left(\frac{dy}{dx}\right) \Delta x$.

This is the *small increment formula* (also called the *linear approximation formula*), relating a small change Δy in y to a small change Δx in x through the derivative.

It essentially approximates the curve by its tangent line at that point.

For example, if $y = x^3$, then $\frac{dy}{dx} = 3x^2$, so the small increment formula gives $\Delta y \approx (3x^2)\Delta x$.

If we take $x = 2$ and $\Delta x = 0.1$, this gives $\Delta y \approx (3)(4)(0.1) = 1.2$.

The exact value is $\Delta y = f(2.1) - f(2) = (2.1)^3 - (2)^3 = 1.261$.

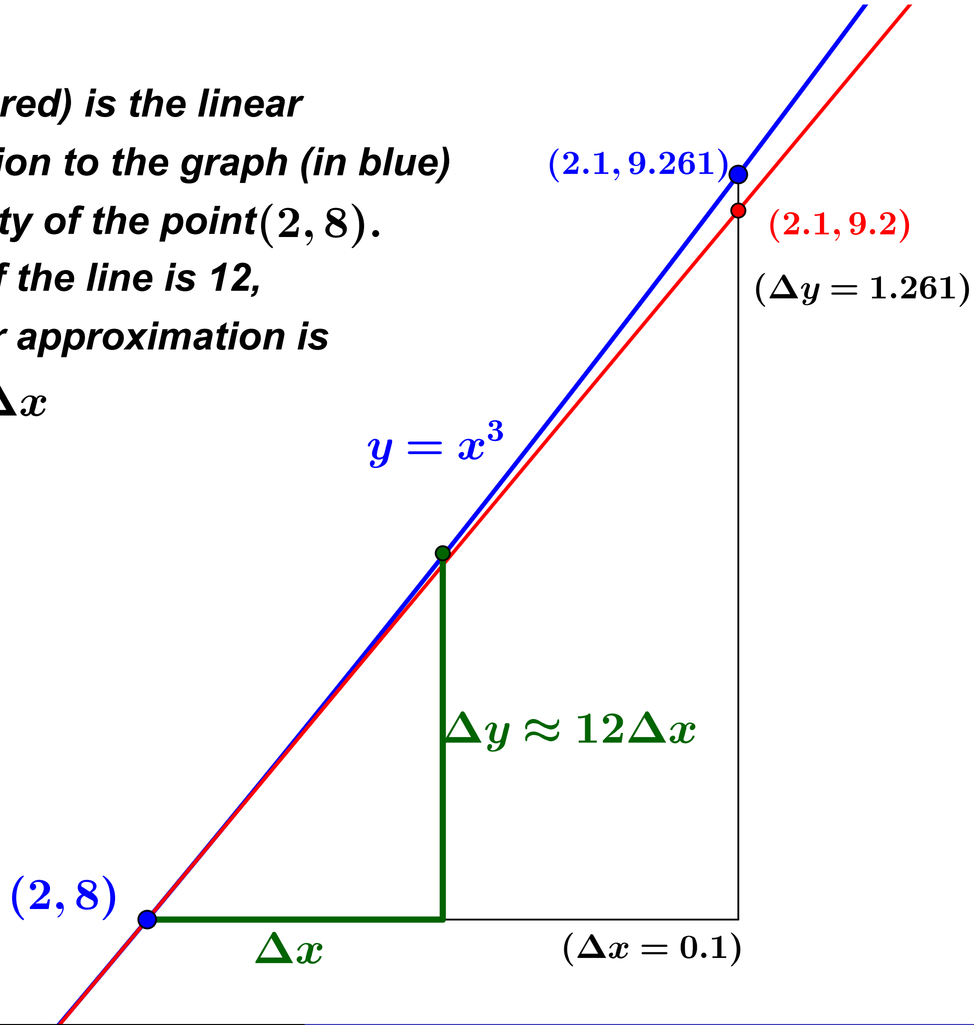
Why should we use an approximation when the exact answer can easily be calculated?

The reason is that we can get a formula for any small increment Δx without having to calculate $f(2 + \Delta x)$ each time. So in the above example with $x = 2$, we get $\Delta y \approx (3)(4)(\Delta x) = 12\Delta x$.

This tells us how y responds to small changes in x near $x = 2$. It is called the *linear approximation to the curve* near $x = 2$.

The line (in red) is the linear approximation to the graph (in blue) in the vicinity of the point* $(2, 8)$. *The slope of the line is 12, so the linear approximation is

$$\Delta y \approx 12\Delta x$$



Essentially, we are using the tangent line to the curve at $x = 2$ as an approximation to the curve itself in the vicinity of $x = 2$.

More generally, when we have a function $y = f(x)$, with derivative $\frac{dy}{dx}$, when we change $x \to x + \Delta x$, then the corresponding change in y is given approximately by $y \to y + \Delta y \approx y + \frac{dy}{dx}\Delta x$.
Again, we are just approximating the function at a point by its tangent line there; its linear approximation.

This is the *first order Taylor expansion* and is often written in terms of $f(x)$ as

$$f(x + \Delta x) \approx f(x) + f'(x)\Delta x.$$

Linear approximation of functions of two variables

Consider a function of two variables $z = f(x, y)$.

How does z vary when we change x and y?

If x changes by a small increment Δx and y is held constant, then the change in z, that is Δz, can be approximated by $\Delta z \approx \frac{\partial z}{\partial x} \Delta x$.

Likewise if y changes by a small increment Δy and x is held constant, then the change in z can be approximated by $\Delta z \approx \frac{\partial z}{\partial y} \Delta y$.

If both change, that is $x \to x + \Delta x$ and $y \to y + \Delta y$, then the change in z can be approximated by

$$\Delta z \approx \frac{\partial z}{\partial x} \Delta x + \frac{\partial z}{\partial y} \Delta y.$$

This is called the *linear approximation formula* for functions of two variables.

Recall that Δz is the change in z, that is

$\Delta z = z(x + \Delta x, y + \Delta y) - z(x, y)$.

So $z(x + \Delta x, y + \Delta y) \approx z(x, y) + \frac{\partial z}{\partial x} \Delta x + \frac{\partial z}{\partial y} \Delta y$.

This is called the *first order Taylor expansion* of z at (x, y).

Example

Consider the function $z = x^3y - xy^3$*.*

(i) Find the partial derivatives of z *at the point* $(2, 1)$*.*

(ii) Write out the first order Taylor expansion of z *near* $(2, 1)$*.*

(iii) Use this to estimate $z(2.1, 0.8)$ *and compare with the exact value.*

(i) The first partial derivatives are $\dfrac{\partial z}{\partial x} = 3x^2y - y^3$ and $\dfrac{\partial z}{\partial y} = x^3 - 3xy^2$.

When $(x, y) = (2, 1)$, we get $\dfrac{\partial z}{\partial x} = 3(2)^2(1) - (1)^3 = 11$

and $\dfrac{\partial z}{\partial y} = 2^3 - 3(2)(1)^2 = 2$.

(ii) The first order Taylor expansion is then

$$z(2 + \Delta x, 1 + \Delta y) \approx z(2, 1) + \frac{\partial z}{\partial x}\Delta x + \frac{\partial z}{\partial y}\Delta y$$

$$= 6 + 11\Delta x + 2\Delta y.$$

(Note we used $z(2, 1) = 2^3(1) - 2(1^3) = 6$.)

(iii) To get $z(2.1, 0.8)$, we put $\Delta x = 0.1$ and $\Delta y = -0.2$.

So $z(2.1, 0.8) \approx 6 + 11(0.1) + 2(-0.2) = 6.7$.

The exact value is

$$z(2.1, 0.8) = (2.1)^3(0.8) - 2.1(0.8)^3 = 6.3336.$$

General Comments

1. The error in the first order Taylor expansion involves terms such as $(\Delta x)^2$, $(\Delta y)^2$, $\Delta x \Delta y$ (called second order terms) and higher powers.

So we write the exact equation

$$z(x + \Delta x, y + \Delta y) = z(x, y) + \frac{\partial z}{\partial x}\Delta x + \frac{\partial z}{\partial y}\Delta y + O(2),$$

where $O(2)$ includes all second order and higher terms.

Sometimes, for increase accuracy, second order terms are included.

The accuracy of all these approximations improves as Δx and Δy get smaller.

2. Sometimes for brevity, we use the z_x notation for partial derivatives and write the Taylor expansion as

$$\Delta z \approx z_x \Delta x + z_y \Delta y$$

or putting in the (x, y) dependence

$$z(x + \Delta x, y + \Delta y) \approx z(x, y) + z_x(x, y)\Delta x + z_y(x, y)\Delta y.$$

We can extend this this to any number of independent variables. Say $z = z(u, v.w)$.

Then the first order Taylor expansion is

$$\Delta z \approx z_u \Delta u + z_v \Delta v + z_w \Delta w.$$

Examples

1. *A* 10×10 *square metal plate is described by coordinates* $0 \leqslant x \leqslant 10, \quad 0 \leqslant y \leqslant 10$*. The temperature at any point* (x, y) *of the plate is given by*

$$T(x, y) = 30 + 2x - \frac{y^2}{50} + \frac{x^4}{100}.$$

Find an approximation to T *at points near* $(2, 5)$*.*

The first order Taylor expansion near $(2, 5)$ is

$$T(2 + \Delta x, 5 + \Delta y) \approx T(2, 5) + T_x(2, 5)\Delta x + T_y(2, 5)\Delta y.$$

First we get the partial derivatives

$$T_x(x, y) = 2 + \frac{4x^3}{100} \quad \text{and} \quad T_y(x, y) = -\frac{2y}{50}.$$

Now we evaluate all these at $(2, 5)$.

First $T(2, 5) = 30 + 2(2) - \dfrac{25}{50} + \dfrac{16}{100} = 33.66.$

Next $T_x(2, 5) = 2 + \dfrac{(4)(8)}{100} = 2.32.$

and $T_y(2, 5) = -\dfrac{(2)(5)}{50} = -0.2.$

Plugging these values in then gives

$$T(2 + \Delta x, 5 + \Delta y) \approx 33.66 + 2.32\Delta x - 0.2\Delta y.$$

This tell us that as we move from $(2, 5)$, the temperature rises quickly with increasing x and falls slowly with increasing y.

2. *The pH of a solution containing three dissolved substances, with concentrations* $x_1, x_2, x_3,$ *is*

$$\phi(x_1, x_2, x_3) = (7 + x_2)\, e^{0.1x_1 - 0.5x_3}$$

Assume that initially, $x_1 = 5,\ x_2 = 2,\ x_3 = 1.$ *Investigate how* ϕ *responds to small changes in* x_1, x_2 *and* x_3*.*

The Taylor approximation is

$$\Delta\phi \approx \phi_{x_1}\Delta x_1 + \phi_{x_2}\Delta x_2 + \phi_{x_3}\Delta x_3.$$

We need to find the three partial derivatives:

$$\phi_{x_1} = 0.1(7 + x_2)e^{0.1x_1 - 0.5x_3},$$

$$\phi_{x_2} = e^{0.1x_1 - 0.5x_3},$$

$$\phi_{x_3} = -0.5(7 + x_2)e^{0.1x_1 - 0.5x_3}.$$

Substituting in the values $x_1 = 5,\ x_2 = 2,\ x_3 = 1,$ and noting that $e^{0.1x_1 - 0.5x_3} = e^0 = 1$, we get

$$\phi(5, 2, 1) = 9, \qquad \phi_{x_1}(5, 2, 1) = 0.9,$$

$$\phi_{x_2}(5, 2, 1) = 1, \qquad \phi_{x_3}(5, 2, 1) = -4.5.$$

So the Taylor approximation is

$$\phi(5 + \Delta x_1, 2 + \Delta x_2, 1 + \Delta x_3) \approx 9 + 0.9\Delta x_1 + \Delta x_2 - 4.5\Delta x_3.$$

This shows that the pH of the solution is most sensitive to changes in x_3, which decreases acidity as its concentration increases.

Relative Change and Percentage Change

Suppose a quantity x changes by $\Delta x = 5$. If $x = 10$, this is a big change, whereas if $x = 1000$, this is a small change. So we need the concept of *relative change*.

We define the *relative change* to be $\frac{\Delta x}{x}$.

The *percentage change* is just the relative change expressed as a percentage.

So if $x = 10$ and $\Delta x = 5$,
the relative change is $\frac{\Delta x}{x} = \frac{5}{10} = 0.5 = 50\%$.
If $x = 1000$ and $\Delta x = 5$,
the relative change is $\frac{\Delta x}{x} = \frac{5}{1000} = 0.005 = 0.5\%$.

Analysis of Measurement (Observational) Error

When a quantity is measured in an experiment, errors can arise in two ways:

- **Random errors** These are caused by random effects, such as weather, human error *etc.* They are analysed using statistics. When the experiment is repeated, different random errors may occur.
- **Systematic errors** These are errors that are reproduced when the experiment is repeated, often due to inaccurate or faulty instruments.

We will focus on random errors in measurement that give rise to uncertainty in the result of an experiment.

Suppose we measure a quantity and get the answer x.

Suppose the true value, which may be unknown, is x_{true}.

Then we write $x = x_{true} + \delta x$,

where the error $\delta x =$ measured value $-$ true value.

The error δx is random and may be positive or negative.

The quantity $|\delta x|$ is called the **absolute error**.

We define the **relative error** to be $e_x = \left|\frac{\delta x}{x}\right|$,

and the **percentage error** to be $100e_x\%$.

(*e.g.* a relative error of $e_x = 0.05$ can be written as $e_x = 5\%$.)

We will specify measurement error in one of two ways:

- **Maximum Absolute Error** This says the error (in absolute value) is no more than some specified value. For example, we might say $|\delta x| \leqslant 1$mm. Essentially this says δx is uniformly distributed with $-1 \leqslant \delta x \leqslant 1$.
- **Measurement Uncertainty/Standard Deviation** This says δx is a random variable (often taken to be normally distributed), with a specified standard deviation.

Measurement Uncertainty In an experiment, a measurement is often stated with a given **uncertainty**, *e.g.* a length might given as $x = 128\text{mm} \pm 1\text{mm}$. This says that x is a random variable with estimated mean 128mm and standard deviation 1mm.

More generally, we denote the *standard deviation* of x (and hence of δx) by σ_x.

In the case above, $\sigma_x = 1$mm; if we assume the errors are normally distributed, we could say $|\delta x| \leqslant 1$mm with probability 68% or $|\delta x| \leqslant 2$mm with probability 95%.

Unlike maximum absolute error, which can never be exceeded, values of say $x + 2\sigma_x$ are achievable, but with low probability. (For example, $x = (128 + 2)$mm.)

Error propagation (single variable)

Suppose we measure a quantity to be x with error δx.

The relative error is $e_x = \dfrac{|\delta x|}{|x|}$.

Suppose we have some other quantity derived from x through the function $y = f(x)$. How is δy, the error in y, related to δx? How is the relative error in e_y related to e_x?

We use the first order linear approximation $\delta y \approx \dfrac{dy}{dx}\delta x$.

So the absolute error in y is $|\delta y| \approx \left|\dfrac{dy}{dx}\delta x\right|$ and the relative error is $e_y = \dfrac{|\delta y|}{|y|} = \dfrac{1}{|y|}\left|\dfrac{dy}{dx}\right||\delta x| = \dfrac{|x|}{|y|}\left|\dfrac{dy}{dx}\right| e_x$.

For example, if $y = x^n$, then $\dfrac{dy}{dx} = nx^{n-1}$,

so $\dfrac{|x|}{|y|}\left|\dfrac{dy}{dx}\right| = \dfrac{|x|}{|x^n|}|nx^{n-1}| = \dfrac{|nx^n|}{|x^n|} = |n|$.

So $e_y = |n|e_x$, that is the relative error in x^n is $|n|$ times the relative error in x!

For example, if $y = x^3$, then $e_y = 3e_x$.

Or if $y = \dfrac{1}{x} = x^{-1}$, then $e_y = |-1|\, e_x = e_x$.

Note that since $(\sigma_x)^2 = \mathrm{Var}(\delta x)$ and $(\sigma_y)^2 = \mathrm{Var}(\delta y)$,

it follows that the standard deviations are connected by

$\sigma_y = \left|\dfrac{dy}{dx}\right| \sigma_x$.

Error propagation (several variables)

We just look at a function of two variables $z = f(x, y)$.

Suppose we have measurements of x and y (which we assume are uncorrelated), with errors δx and δy.

Maximum Absolute and Relative Errors

Suppose we are given the maximum absolute errors in x and y, namely $|\delta x| \leqslant \delta x_{max}$ and $|\delta y| \leqslant \delta y_{max}$.

How do we find the maximum absolute error and the maximum relative error in z?

We start with the linear approximation $\delta z = z_x \delta x + z_y \delta y$.

This gives the error in z in terms of the errors in x and y.

Taking absolute values, and using the Triangle Inequality, we get

$$|\delta z| = |z_x \delta x + z_y \delta y| \leqslant |z_x|\,|\delta x| + |z_y|\,|\delta y| \leqslant |z_x|\,\delta x_{max} + |z_y|\,\delta y_{max}.$$

So the maximum absolute error is

$$\delta z_{max} = |z_x|\,\delta x_{max} + |z_y|\,\delta y_{max}.$$

The maximum relative errors in x and y and z are

$(e_x)_{max} = \dfrac{\delta x_{max}}{|x|}$ and $(e_y)_{max} = \dfrac{\delta y_{max}}{|y|}$

and $(e_z)_{max} = \dfrac{\delta z_{max}}{|z|}$.

So $\delta x_{max} = |x|\,(e_x)_{max}$ and $\delta y_{max} = |y|\,(e_y)_{max}$ and $\delta z_{max} = |z|\,(e_z)_{max}$.

So $|z|\,(e_z)_{max} = |z_x|\,|x|\,(e_x)_{max} + |z_y|\,|y|\,(e_y)_{max}$.

That is, $(e_z)_{max} = \dfrac{|xz_x|}{|z|}(e_x)_{max} + \dfrac{|yz_y|}{|z|}(e_y)_{max}$.

This is the *maximum relative error* in z, when the maximum relative errors in x and y are $(e_x)_{max}$ and $(e_y)_{max}$.

Measurement Uncertainty (Standard Deviation)

Suppose the measurements x and y have uncertainties (standard deviations) σ_x and σ_y.

(Note: these are also the standard deviations of the errors δx and δy.)

We want the uncertainty (standard deviation) σ_z in z.

We go back to the error formula $\delta z = z_x \delta x + z_y \delta y$.

A result in statistics states that if X and Y are uncorrelated random variables, then
$\mathrm{Var}(aX + bY) = a^2\mathrm{Var}(X) + b^2\mathrm{Var}(Y)$.

Applying this, we get

$\mathrm{Var}(\delta z) = \mathrm{Var}(z_x\delta x + z_y\delta y) = z_x^2\mathrm{Var}(\delta x) + z_y^2\mathrm{Var}(\delta y)$.

That is, $\sigma_z^2 = z_x^2\sigma_x^2 + z_y^2\sigma_y^2$.

This gives the uncertainty in z as:

$$\sigma_z = \sqrt{(z_x)^2\sigma_x^2 + (z_y)^2\sigma_y^2} = \sqrt{\left(\frac{\partial z}{\partial x}\right)^2\sigma_x^2 + \left(\frac{\partial z}{\partial y}\right)^2\sigma_y^2}.$$

This is the well-known root-mean-square (RMS) formula in error propagation.

Examples

1(a). *The volume of a cylinder is* $V = \pi r^2 h$. *It desired to manufacture a cylinder of radius* $r = 10$ *and height* $h = 30$. *It is known that the relative error in the measurement of* r *is at most* 2% *and in* h *is at most* 1%. *Calculate the maximum absolute error and maximum percentage error in* V.

What we are told is that $e_r \leqslant 0.02$ and $e_h \leqslant 0.01$.

So $|\delta r| = re_r \leqslant (10)(0.02) = 0.2$,

and $|\delta h| = he_h \leqslant (30)(0.01) = 0.3$.

The first order linear approximation says that

$\delta V \approx V_r \delta r + V_h \delta h$.

The partial derivatives are $V_r = \dfrac{\partial(\pi r^2 h)}{\partial r} = 2\pi r h$ and $V_h = \dfrac{\partial(\pi r^2 h)}{\partial h} = \pi r^2$.

So $\delta V \approx 2\pi r h\, \delta r + \pi r^2\, \delta h = 600\pi \delta r + 100\pi \delta h$.

So $|\delta V|$

$\leqslant 600\pi|\delta r| + 100\pi|\delta h| \leqslant 600\pi(0.2) + 100\pi(0.3) = 150\pi$.

So the maximum absolute error $\delta V_{max} = 150\pi$.

The volume of the desired cylinder is

$V = \pi r^2 h = \pi(10)^2(30) = 3000\pi$.

So the maximum relative error is

$$(e_V)_{max} = \frac{\delta V_{max}}{|V|} = \frac{150\pi}{3000\pi} = 0.05 = 5\%.$$

A quick way to do this is to note that the maximum relative error in r^2 is twice the maximum relative error in r, that is 4% and the maximum relative error in h is 1% so the maximum relative error in $V = \pi r^2 h$ is $4\% + 1\% = 5\%$.

1(b). *We have the same cylinder as in* **1(a).**
We are given that $r = 10 \pm 0.2$ *and* $h = 30 \pm 0.3$.
Find the uncertainty in V.

What we are told is that the uncertainties in r and h have standard deviations $\sigma_r = 0.2$ and $\sigma_h = 0.3$.

So, by the RMS formula, the uncertainty in V has standard deviation $\sigma_V = \sqrt{(V_r)^2\sigma_r^2 + (V_h)^2\sigma_h^2}$.

As above, $V_r = 2\pi rh = 2\pi(10)(20) = 600\pi$ and $V_h = \pi r^2 = 100\pi$.

So $\sigma_V = \sqrt{(600\pi)^2(0.2)^2 + (100\pi)^2(0.3)^2}$
$= \sqrt{(120\pi)^2 + (30\pi)^2} = \sqrt{(16+1)(30\pi)^2} = 30\pi\sqrt{17}$.

Comment: This is smaller than the maximum absolute error of 150π above (since $\sqrt{17} < 5$), but values bigger than this may be achieved.

2. *In the chemical reaction* $X + 2Y \rightleftharpoons 3Z$,
the equilibrium coefficient for the reaction is defined to be
$$K = \frac{z^3}{xy^2} = z^3x^{-1}y^{-2},$$
where x, y, z *are the concentrations of* X, Y, Z *respectively.*

Assume $(e_x)_{max} = (e_y)_{max} = (e_z)_{max} = 1\%$.

Find the maximum percentage error $(e_K)_{max}$ *in* K.

Here $K = z^3x^{-1}y^{-2}$ is a function of three variables.

Extending the relative error formula to three variables gives:

$$e_K \leqslant \frac{|xK_x|}{|K|}e_x + \frac{|yK_y|}{|K|}e_y + \frac{|zK_z|}{|K|}e_z.$$

First $\dfrac{|xK_x|}{|K|} = \dfrac{|x(-1)(z^3x^{-2}y^{-2})|}{|z^3x^{-1}y^{-2}|} = 1.$

Next $\dfrac{|yK_y|}{|K|} = \dfrac{|y(-2)(z^3x^{-1}y^{-3})|}{|z^3x^{-1}y^{-2}|} = 2.$

Next $\dfrac{|zK_z|}{|K|} = \dfrac{|z(3)(z^2x^{-1}y^{-2})|}{|z^3x^{-1}y^{-2}|} = 3.$

So $e_K \leqslant (1)e_x + (2)e_y + (3)e_z \leqslant 0.01 + 0.02 + 0.03 = 0.06.$

So the maximum percentage error in K is $(e_K)_{max} = 0.06 = 6\%$.

Comments

- Each power just gives the absolute value of the power times the relative error.
  This follows since $\dfrac{\delta(x^n)}{x^n} = \dfrac{nx^{n-1}\delta x}{x^n} = n\dfrac{\delta x}{x}$.
  So $e_{x^n} = |n|e_x$. That is the relative error in x^n is n times the relative error in x.
- Positive and negative powers have the same effect on relative error.
- When quantities are multiplied or divided, the combined relative error just adds the individual relative errors.

So for example, if we had $K = x^3y^{-4}z^5$ instead, then we'd get $(e_K)_{max} = (3)(e_x)_{max} + (4)(e_y)_{max} + (5)(e_z)_{max}$.