MATH 442: Mathematical Modeling

Instructor: Prof. Wolfgang Bangerth
bangerth@math.tamu.edu

Answers to the midterm exam 10/21/2010

Problem 1 (Taylor expansion).
Part a: Given the function of one argument $f(x) = \sin 2x$, find its Taylor expansion around $x_0 = 0$ up to terms that are cubic in x. **(8 points)**
Part b: Given the function of two arguments $f(x, y) = (\sin x)e^{2y}$, find its Taylor expansion around $x_0 = 0, y_0 = 1$ up to terms that are linear in x and y individually (i.e. including the term that is proportional to $x \cdot y$). **(8 points)**
Part c (in words): In general, describe what happens if we plotted a function $f(x)$ and its Taylor expansions to higher and higher orders? Is your observation true for all functions $f(x)$ or only if $f(x)$ has certain properties? If so, which properties? **(12 points)**

Answers.
Part a: The general formula for the Taylor expansion around x_0 of a function of one argument is

$$f(x) \approx \frac{1}{0!}f(x_0) + \frac{1}{1!}f'(x_0)(x - x_0) + \frac{1}{2!}f''(x_0)(x - x_0)^2 + \frac{1}{3!}f'''(x_0)(x - x_0)^3 + \dots.$$

Derivatives are easily computed and evaluated at x_0 here:

$$\begin{aligned} f(x) &= \sin 2x & f(x_0) &= \sin 0 = 0 \\ f'(x) &= 2\cos 2x & f'(x_0) &= 2\cos 0 = 2 \\ f''(x) &= -4\sin 2x & f''(x_0) &= -4\sin 0 = 0 \\ f'''(x) &= -8\cos 2x & f'''(x_0) &= -8\cos 0 = -8. \end{aligned}$$

The solution is then

$$\sin(x) \approx 0 + 2x + 0 - \frac{8}{6}x^3$$

Part b: The correct general formula for a function of two arguments looks like this (check the Wikipedia page and your multivariate calculus textbooks for more background):

$$f(x, y) \approx \frac{1}{0!}f(x_0, y_0) + \frac{1}{1!}\frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{1}{1!}\frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) + 2\frac{1}{2!}\frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(x - x_0)(y - y_0) + \dots.$$

Note that the last term is still linear in x and y separately, even though it involves second derivatives of f.

For the given function and $x_0 = 0, y_0 = 1, we have$

$$\begin{aligned}
f(x_0, y_0) &= \sin x_0 e^{2y_0} = 0 \\
\frac{\partial f}{\partial x}(x_0, y_0) &= \cos x_0 e^{2y_0} = e^2 \\
\frac{\partial f}{\partial y}(x_0, y_0) &= 2 \sin x_0 e^{2y_0} = 0 \\
\frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) &= 2 \cos x_0 e^{2y_0} = 2e^2.
\end{aligned}$$

The final result is therefore

$$\sin x e^{2y} \approx e^2 x + 2e^2 x(y-1).$$

There were many answers that considered terms like $f'(x, y) = \cos xe^{2y} + 2 \sin xe^{2y}$ or $f'(x, y) = 2 \cos xe^{2y}$. This does not make any sense. Think of $f(x, y)$ as the function that describes the height of a surface at a point x, y. Partial derivatives are then a rate of change, i.e. how does the height change as we move in x or y direction. In other words, speaking of $f'(x, y)$ is meaningless: what do you mean by *the* derivative of f – its x derivative or its y derivative? Without answering this question it also doesn't make sense to compute the derivative as $\cos xe^{2y} + 2 \sin xe^{2y}$ or $2 \cos xe^{2y}$.

Part c: As you plot the original function together with its Taylor expansions containing more and more terms, the Taylor expansions will come closer and closer to the original function. This is only true, however, if the function is continuously differentiable to at least the order of the Taylor approximation on the interval on which you plot the two together. If you want to repeat the process infinitely often, you need that the the function is analytic (which is roughly, but not quite, equivalent to saying that it is infinitely often continuously differentiable).

It is important to realize these restrictions. For example, if we take the function $f(x) = e^x$ (a function that is analytic) you can get as close as you want with your Taylor expansion. But this is not so if the function is not continuous. For example, consider the function that is zero for $x < 0$ and one for $x \geq 0$: if we tried to find its Taylor expansion around $x_0 = 1$ we have that

$$\begin{aligned}
f(x) &\approx \frac{1}{0!} f(x_0) + \frac{1}{1!} f'(x_0)(x - x_0) + \frac{1}{2!} f''(x_0)(x - x_0)^2 + \dots \\
&= 1 + 0(x-1) + 0(x-1)^2 + \dots \\
&= 1
\end{aligned}$$

because all derivatives of this function are zero at $x_0 = 1$. The Taylor series is then a good approximation for positive x (where $f(x) = 1$), but a bad approximation for $x < 0$ – and the Taylor series does not become a better approximation the more terms we take into account (they are all zero after all).

You can play similar games with higher derivatives. For example, the function $f(x) = |x|$ has a kink at $x = 0$ but it is at least continuous at this point. If we try to find its Taylor approximation at $x_0 = 1$ we find

$$
\begin{aligned}
f(x) &\approx \frac{1}{0!}f(x_0) + \frac{1}{1!}f'(x_0)(x - x_0) + \frac{1}{2!}f''(x_0)(x - x_0)^2 + \ldots \\
&= 1 + 1(x - 1) + 0(x - 1)^2 + \ldots \\
&= x
\end{aligned}
$$

Again, this us a good approximation for $x > 0$ but since the Taylor series at $x_0 = 1$ does not "see" the kink at $x = 0$ it is a bad approximation for $x < 0$ and the approximation does again not become better as we take more and more terms into account. You can play similar games with functions that are continuous but only once continuously differentiable at a point (e.g. $f(x) = x\,|x|$), continuous but only twice continuously differentiable at a point (e.g. $f(x) = x^2\,|x|$), etc.: Their Taylor approximation does not become better and better as we take more and more terms into account.

Problem 2 (Dimensional analysis and non-dimensionalization.) The trajectory of an electron in a magnetic field is given by Newton's law $\vec{F}(t) = m\ddot{\vec{x}}(t)$ where the force is the Lorentz force, $\vec{F}(t) = e\dot{\vec{x}}(t) \times \vec{B}$. Consequently, the equations of motion are

$$m\ddot{\vec{x}}(t) = e\dot{\vec{x}}(t) \times \vec{B}.$$

Here, m is the mass of an electron measured in kilograms, e is its charge measured in Coulombs, and $\vec{B}$ is the magnetic field measured in Teslas (1 Tesla equals 1 kilogram per Coulomb per second, $kg\ C^{-1}\ s^{-1}$).
Part a: Find a typical time scale T for this system by appropriate combination of the various coefficients above. For this, you may wish to write the magnetic field vector as $\vec{B} = B_0\vec{\beta}$ where B_0 is the magnetic field *strength* (still measured in Teslas) and $\vec{\beta}$ is the *direction* of the field (i.e. it is a vector with unit magnitude and no physical dimensions). **(12 points)**
Part b: By measuring time in multiples of the time scale T, derive equations that no longer contain any coefficients and only involve $x(t)$ and the unit direction $\vec{\beta}$. **(12 points)**

Answers.

Part a: The only coefficients with physical units you have in this problem are m, e, B_0. The only way to combine them into something that has units seconds is

$$T = \frac{m}{eB_0}.$$

I have seen a number of answers that considered expressions like $\frac{m}{e\vec{B}}$, $\frac{\ddot{\vec{x}}}{\dot{\vec{x}}}$, or similar. Independently how you got there you should have realized that this can't be correct: what does it mean if you have a fraction where you have a vector in the denominator?

Part b: With this time scale T, let us change variables as $t = \hat{t}T$, i.e. we want to measure times not in seconds any more but in terms of (unitless) multiples $\hat{t}$ of the time scale T. Re-arranging the original equation a bit yields

$$\underbrace{\frac{m}{eB_0}}_{=T} \frac{d}{dt}\left(\frac{d}{dt}\vec{x}(t)\right) = \frac{d}{dt}\vec{x}(t) \times \vec{\beta}.$$

Now, let's consider these derivatives. We have by a simple change of variables that

$$\frac{d}{dt}\vec{x}(\hat{t}T) = \frac{d\hat{t}}{dt}\frac{d}{d\hat{t}}\vec{x}(\hat{t}T)$$

and because $\hat{t} = \frac{t}{T}$:

$$\frac{d}{dt}\vec{x}(\hat{t}T) = \frac{1}{T}\frac{d}{d\hat{t}}\vec{x}(\hat{t}T).$$

Similarly for the second time derivatives. This yields for the equation in question:

$$T\frac{1}{T^2}\frac{d^2}{d\hat{t}^2}\vec{x}(\hat{t}T) = \frac{1}{T}\frac{d}{d\hat{t}}\vec{x}(\hat{t}T) \times \vec{\beta}.$$

After canceling factors of T all that remains is

$$\frac{d^2}{d\hat{t}^2}\vec{x}(\hat{t}T) = \frac{d}{d\hat{t}}\vec{x}(\hat{t}T) \times \vec{\beta}.$$

If you want, you can introduce a new variable $\vec{X}(\hat{t}) = \vec{x}(\hat{t}T)$ which then has to satisfy the equation

$$\frac{d^2}{d\hat{t}^2}\vec{X}(\hat{t}) = \frac{d}{d\hat{t}}\vec{X}(\hat{t}) \times \vec{\beta}.$$

Solutions to this equation (i.e. trajectories of particles) are circles in planes perpendicular to the direction β.

Problem 3 (Population models with a carrying capacity.) In class and in homework we have considered models of population growth where growth was limited by a finite resource. This led to models like the logistic equation or the Gompertz equation.

Now consider an ecosystem in which one species lives that requires *two* resources. In the spirit of upcoming Halloween you may think, for example, of vampires that require both blood donors as well as coffins; alternatively think of trees that need both light and water. Assume that both necessary resources are finite and that the amount of resource available is independent of the species' population (e.g.: the amount of light and water is fixed and independent of the number of trees). Derive a differential equation that could describe the growth or decline of the population of this species. Describe in words what the various terms and coefficients in your equation mean and how you arrived at them. Also describe in words what characteristics you expect from the solution of your differential equation (e.g., will it grow forever? What happens if the initial number of individuals is very small? If it is very large? If one resource is plentiful and the other is not?). **(24 points)**

Answers. The equation we are looking for needs to satisfy a general principle: the species needs two resources to live so if the number of individuals is so that one of these two resources is exhausted then the overall growth rate must be zero, even if the other resource is still plentiful. In other words, the resource of which less is available presents a limit for population growth. Let us try to model this situation by saying that K_1, K_2 are the amounts of resource that are available to the species, and that c_1, c_2 are the amounts of these two resources that each individual needs. Let $p(t)$ be the population size then the population as a whole consumes fractions $\frac{c_1 p(t)}{K_1}, \frac{c_2 p(t)}{K_2}$ of the resources. The amount that is left over is then $1 - \frac{c_1 p(t)}{K_1}$ and $1 - \frac{c_2 p(t)}{K_2}$. If we let A be the natural growth rate (i.e. the difference between natural births and deaths) then, following the derivation of the logistic model, we could propose this model for population growth:

$$\frac{d}{dt}p(t) = A\left(1 - \frac{c_1 p(t)}{K_1}\right)\left(1 - \frac{c_2 p(t)}{K_2}\right)p(t).$$

This model satisfies our initial requirement: if the population is so that they completely consume one of the resources, i.e. either $c_1 p = K_1$ or $c_2 p = K_2$, then the right hand side becomes zero and the population stops to grow.

I would have given bonus points if you had found out that the model above actually has a problem. Namely, assume for example that the population requires more than there is of resource 1. I.e. $\frac{c_1 p(t)}{K_1} > 1$ and consequently the right hand side becomes negative. This makes sense: if there isn't enough of resource 1, then the population must shrink. The same is true if the population is so large that it consumes more than there is of resource 2. But what if the population is so large that it wants to consume more than there is of *both* resources? In that case, both of the terms in parentheses would be negative and the right hand side as a whole would be positive again, indicating positive population growth. This is clearly not what we had in mind. A better model may therefore propose that the growth rate is solely determined by the resource that is less plentiful, leading to a model of the kind

$$\frac{d}{dt}p(t) = A \min\left\{ \left(1 - \frac{c_1 p(t)}{K_1}\right), \left(1 - \frac{c_2 p(t)}{K_2}\right) \right\} p(t).$$

We could write this equivalently as

$$\frac{d}{dt}p(t) = A \left(1 - \max\left\{ \frac{c_1}{K_1}, \frac{c_2}{K_2} \right\} p(t)\right) p(t),$$

which makes it clear that (i) despite the minimum/maximum operation the model is actually differentiable (because c_1, K_1, c_2, K_2 are just constant, so the maximum is just a number), (ii) the model simply ignores the more abundant resource. Other variants of these equations are of course also possible.

Let me also comment on one model that would not be correct. Consider for example

$$\frac{d}{dt}p(t) = A \left(1 - \frac{c_1 p(t)}{K_1} - \frac{c_2 p(t)}{K_2}\right) p(t).$$

This model does not quite represent what we want: the growth rate would be zero if the population as a whole consumed only half of the two resources, at which point we would still expect it to grow. Similarly, a model of the kind

$$\begin{aligned}\frac{d}{dt}p(t) &= A \left[\left(1 - \frac{c_1 p(t)}{K_1}\right) + \left(1 - \frac{c_2 p(t)}{K_2}\right) \right] p(t) \\ &= A \left(2 - \frac{c_1 p(t)}{K_1} - \frac{c_2 p(t)}{K_2}\right) p(t)\end{aligned}$$

would not work: if the population consumes *all* of one resource but only a small part of the other, then the term in square brackets would still be positive, indicating growth.

Qualitatively, what we expect from solutions to any correct model is this: If the initial population is small then it will grow quickly until the resource that is less plentiful limits growth and makes the population number level out. If we start with a large population, larger than the carrying capacity of the less abundant resource, then the population will decline until it hits the carrying capacity of the less available resource.

Note, finally, that the carrying capacity is that population for which overall growth is zero. In other words, the carrying capacity equals

$$\min\left\{ \frac{K_1}{c_1}, \frac{K_2}{c_2} \right\}.$$

In other words, the carrying capacity of the ecosystem equals the lesser of the ratio of the amounts of two resources available divided by the amount each individual of the species needs of it.

Problem 4 (Fitting parameters.) Let's say you have reason to believe that the concentration of a substance in a chemical reactor grows over time as $y(t) = K(1 - e^{-rt})$. You don't know the values of the coefficients K, r but you have taken measurements: at time $t_i, i = 1 \dots N$ you have measured the concentration to be y_i.
Part a: Describe in words and formulas the least squares procedure that can be used to find best-fit parameters K, r given the data t_i, y_i. **(12 points)**
Part b: If you had M possible models $y_k(t), k = 1 \dots M$ (the function $y(t)$ above could be one of them) and each model has unknown parameters, describe in words and formulas a criterion that can be used to compare which of the models best describes the data. Recall that you did this when you compared how well the linear and exponential models described the population growth in the United States over the past century. **(12 points)**

Answers.
Part a: To find the best fit can be found by minimizing the sum of squared differences between our data points and the predicted model curve. Let us define the misfit as

$$m(r, K) = \sum_{i=1}^{N} \left(y_i - K(1 - e^{-rt})\right)^2$$

then the best parameters minimize $m(r, K)$. A necessary condition for a minimizer is that

$$\frac{\partial m(r, K)}{\partial r} = 0, \qquad \frac{\partial m(r, K)}{\partial K} = 0.$$

Part b: If r^*, K^* are the parameters that minimize $m(r, K)$ (in other words, $m(r^*, K^*) \leq m(r, K)$ for all other values of r, K), then we can compute a value $m^* := m(r^*, K^*)$ that indicates how far the best fit curve is away from the data points.

We can compute this sort of measure for all the K models we have. Let's say we call them $m_1^*, \dots, m_K^*$. Each of these numbers tells us how far the corresponding curve is away from the data points. The best model is then the one for with the least value m_k^*.

In one of the homeworks you were asked to compute the root mean square error (RMSE) for the best fit models. For the model in the first part of this question, the formula for the RMSE is

$$RMSE(r, K) = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y_i - K(1 - e^{-rt}))^2}.$$

That means, we have $RMSE(r, K) = \sqrt{\frac{1}{N} m(r, K)}$. Since the square root is a monotonous function, that model for which m^* is minimal is also the one for which $RMSE^*$ is minimal. In other words, we can both use the misfit as well as the RMSE to compare how good different models are.