Part I

Principles of mathematical modeling

Motivation

In this part we will go through basic principles of mathematical modeling and their connection to simulations.

Questions that we will address are

- What is a mathematical model?
- What do we mean with a *good* mathematical model?
- What is a numerical model?
- How to classify mathematical models?
- How to use mathematical models and simulations to solve problems?
- Which type of errors should we be aware of when working with mathematical models and simulations?

This part is based on Chapter 1 of the book of Kai Velten, and Chapter 1 of the book of Quarteroni and Gervasio.

What is a model?

First: What is a model?
And: Why do we need a model?

- A model is used to break down the complexity of a system to describe it as simple as possible.
- A model is used to answer a question about the system.

Example 1.1: Model for a bike

- How to cycle?
- Why is the bike not moving forward?

What is a model?

Definition 1.2: Model

A *model* of an object is something one can use to answer questions about the object.

A good model should be simple, but still able to answer the question.

Definition 1.3: The best model

The *best model* is the simplest model that still serves its purpose, that is, which is still complex enough to solve the problem at hand.

"Everything should be made as simple as possible, but not simpler"

What is a model?

A modeling scheme generally has the following steps:

1. Question or problem. What is to be answered or solved?
2. System analysis. What is relevant for the question or problem?
3. Modeling. Put together the information from system analysis.
4. Simulation. Use the model on question.
5. Validation. Is the problem solved?

Example 1.1: Model for a bike

Identify the steps for the two bike models.

Warning 1.4: Importance of validation

If the answer to step 5 is no, we do not have a good model.

Warning 1.5: What is a *system*?

System = whatever we are analyzing. Can be single object or collection of processes. E.g. bike, falling skydiver, temperature in a hot water tank, etc.

What is a model?

The system analysis part is usually the most challenging part of a modeling scheme. This step requires knowledge about the system to be studied and can be time consuming.

For a mathematical model, the system analysis step means to identify the equations and parameter values that are needed.

Example 1.6: Skydiving

If someone jumps out of an airplane, what is the speed of that person after 10 seconds?

System analysis: Falls under gravity. Starts with speed of zero.

$$v = v_0 + a \cdot t$$

Modeling:

$$v = 9.8 \cdot t$$

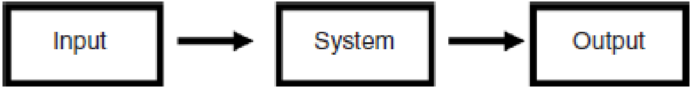
Simulation:

$$v = 9.8 \cdot 10 = 98 \text{ m/s}$$

Validation? If no, what else should be included in the system analysis?

What is a mathematical model?

Mathematical models deal with numbers and mathematical operations (e.g. equations). They are used to answer questions concerning numbers; e.g. how much, how fast, etc.
In the most general form, mathematical models is somehow concerned with input-output systems:



This could be

- An experiment giving different concentrations depending on temperature
- An equation to be solved for different values

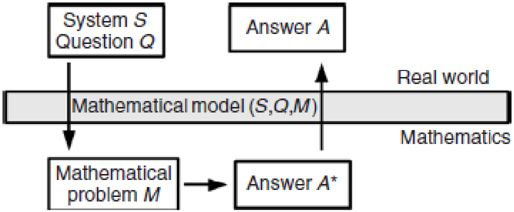
What is a mathematical model?

We here follow the definition of Kai Velten:

Definition 1.7: Mathematical model

A *mathematical model* is a triplet $\{S, Q, M\}$, where S is a system, Q is a question concerning S, and M is a set of mathematical statements which can be used to answer Q.

- System = Whatever we study.
- Question = What we want to find out.
- Mathematical statements = Our path to the answer of the question.



What is a mathematical model?

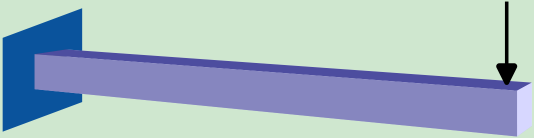
Example 1.8: Height of people in this room

What is the average height of people in this room?

- S: People in this room.
- Q: What is the average height of S?
- M: $\overline{h} = \frac{1}{n} \sum_{i=1}^{n} h_i$

Example 1.9: A bending beam

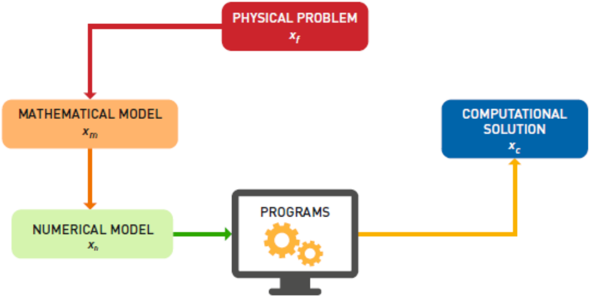
How much will be steal beam of size 10x1x1m attached on one side and under a load of 10 MN on the opposite side bend?



- S: Steal beam of given size.
- Q: How much does it bend under given load?
- M: Partial differential equation (PDE) describing elastic properties of beam.

What is a simulation?

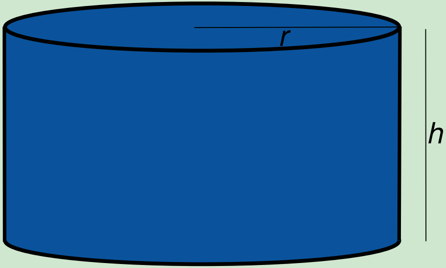
Performing the mathematical statements (or solving them) is the simulation step from slide 5. This might require the use of a computer and the use of a numerical model as well: We perform a *simulation*.



The numerical model is the bridge between the mathematical model and the computer.
In Example 1.8, the simulation is to perform the calculation. In Example 1.9, the simulation is to solve the PDE, which requires a numerical model to find an approximate solution.

Example 1.10: Minimize surface area of hot water tank

Which radius and height should a cylindrical 400 liter hot water tank have to minimize its surface area? (You can assume the walls of the tank have zero width)



- Identify S, Q and M.
- Perform the simulation (find the answer!)

Variables and parameters

We separate between the different terms:

- State variables
  - Describe the state of the system in the mathematical model. Required to answer the question.
- System parameters
  - Describe properties of the system. Needed to compute the state variables.

Example 1.10: Minimize surface area of hot water tank - continued

- State variables: Surface area.
- System parameters: Radius and height.

Variables and parameters

When working with problems with much of information, it can be helpful to early on identify what state variables and system parameters are relevant when setting up the mathematical model.

Example 1.11: Transmitting fluid

Pipe A has diameter 2 cm and can transmit 1000 l/day of water when having a pressure drop of 100 Pa. Pipe B has diameter 3 cm and can transmit 2000 l/day of water with a pressure drop of 120 Pa. The costs of using pipe A and pipe B, is 100 and 400 NOK/day, respectively. If you need to transmit 4000 l and only have 500 NOK available, how can you transmit the water as quick as possible?

- *S*: Two pipes of different capacity and costs.
- *Q*: How to transmit as quick as possible with given budget?
- *M*:
  - State variables: Number of days using pipe A and pipe B: d_A, d_B.
  - System parameters: Prices per day $(100, 400$ NOK/day) and capacities per day $(1000, 2000$ l/day).
  - Not relevant system parameters: Sizes of pipes, pressure drops.

Example 1.11: Transmitting fluid - continued

- M:

$$\begin{aligned} \text{Costs:} \quad & d_A \times 100 + d_B \times 400 \leq 500 \\ \text{Transmitted:} \quad & d_A \times 1000 + d_B \times 2000 = 4000 \end{aligned}$$

Solving second for d_A and insert into first gives

$$d_B \leq 0.5 \text{ while } d_A = 4 - 2d_B$$

To make sure the largest of them is as low as possible, best choice is

$$d_B = 0.5, \quad d_A = 3.$$

$\Rightarrow$ Can transmit the water by using pipe A for three days, and pipe B for half a day.

Variables and parameters

Some advice:

- Check units. What do you get when multiplying numbers with units together?
- Read carefully. Which information is relevant and which is not?

Three general steps for setting up the model:

1. Determine number of unknowns. Usually identified by reading the question to pinpoint what is asked for.
2. Identify all variables and parameters, with units. Give unknowns meaningful symbols.
3. Read problem carefully, translate relevant parts into mathematical statements, making sure units remain meaningful.

Variables and parameters

Example 1.11: Transmitting fluid - continued

1. Number of unknowns: Two, days of using pipe A and days of using pipe B.
2. Unknowns: d_A days, d_B days.
   Prices per day: 100 NOK/day, 400 NOK/day.
   Capacities: 1000 l/day, 2000 l/day.
   Sizes: 2 cm, 3 cm.
   Pressure drops: 100 Pa, 120 Pa.
   Budget: 500 NOK. Amount to transmit: 4000 l.
3. Relevant parts to answer questions are budget and amount to transmit, along with prices and capacities per day.

$$\underbrace{d_A \text{ days} \times 100 \text{ NOK/day}}_{\text{NOK}} + \underbrace{d_B \text{ days} \times 400 \text{ NOK/day}}_{\text{NOK}} \leq 500 \text{ NOK.}$$

$$\underbrace{d_A \text{ days} \times 1000 \text{ l/day}}_{\text{liters}} + \underbrace{d_B \text{ days} \times 2000 \text{ l/day}}_{\text{liters}} \leq 4000 \text{ liters.}$$

Variables and parameters

Sometimes we need to also deal with auxiliary variables. They describe properties of the system which are not needed to answer the question, but are needed for the mathematical formulation of the problem.

Example 1.12: Mixing alloys

The alloys A, B, C contain the following mass fractions (in %) of constituents m, n:

	A	B	C
m	11	7	12
n	14	2	44

If making 1 kg of a new mixture by combining alloys A, B and C, such that 0.5 kg of A is used and the mixture has 100 g of constituent m, how many grams will it contain of constituent n?

- *S*: Three alloys of different composition.
- *Q*: How much of constituent n in new mixture?
- *M*:

Example 1.12: Mixing alloys - continued

1. Number of unknowns: One, grams of n.
2. State variable: M_n g.
   System parameters: Amount of m in A: $11\% = 0.11$ (dimensionless), etc.
   Auxiliary variables: Grams of B and C in new mixture: M_B, M_C.
3. Choose to do everything using grams, to avoid mixing kg and g:

$$\begin{aligned}
\text{Mass of m in new mixture:} \quad & 0.11 \times 500 + 0.07 \times M_B + 0.12 \times M_C = 100 \\
\text{Mass of new mixture:} \quad & 500 + M_B + M_C = 1000 \\
\text{Mass of n in new mixture:} \quad & 0.14 \times 500 + 0.02 \times M_B + 0.44 \times M_C = M_n
\end{aligned}$$

Two first gives $M_B = 300$ g and $M_C = 200$ g, hence the third equation gives us

$$M_n = 164 \text{ g}$$

Warning 1.13: Number of variables and equations

As a rule of thumb: The sum of number of auxiliary variables and state variables should be equal to the number of equations.

Statistical example

Mathematical models can also deal with given data as input-output system.

Example 1.14: Braking of a car

Given the following braking distances of a car under dry conditions[1]:

km/h	m
40	8
50	12.5
60	18
80	32

Assume a linear relation between speed and braking distance. What is the estimated braking distance for 70 km/h?

- S: Braking car.
- Q: When going 70 km/h, what is the expected braking distance?
- M: Linear regression[2], use expression to find estimate for velocity of 70 km/h.

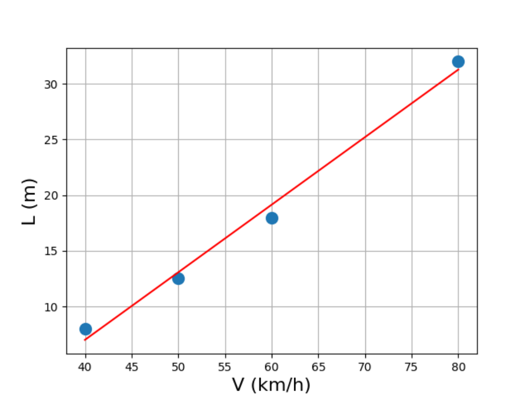
[1] Numbers from Gjensidige

[2] Coming in Part 2

Example 1.14: Braking of a car - continued

Performing the linear regression gives

$$L = 0.6071 \times V - 17.2857$$



For a velocity $V = 70$ km/h, the estimate is $L = 25.2$ m.

Statistical example

Example 1.14 did not require us to know anything about the car itself or how braking works. The given data were sufficient to solve the problem.

The problem could have been considered as a black box. Given the table

Input	Output
40	8
50	12.5
60	18
80	32

and assuming a linear relation between input and output, what is the expected output for an input of 70?

Types of models

Definition 1.15: Phenomenological and mechanistic models

A mathematical model $\{S, Q, M\}$ is called

- *phenomenological* when it is constructed from experimental data only, using no a priori information about S.
- *mechanistic* when statements in M are based on a priori information about S.

Phenomenological models

- are easy to set up.
- are limited by the data in terms of generalizing the model.

Mechanistic models

- require knowledge about the system and are therefore more demanding to set up.
- give a better understanding of the behavior of the system.
- can more easily be generalized.

Types of models

Definition 1.16: Stationary and instationay models

A mathematical model $\{S, Q, M\}$ is called

- *instationary* (or time-dependent), if at least one of its system parameters or state variables depend on time.
- *stationary* otherwise.

Definition 1.17: Distributed and lumped models

A mathematical model $\{S, Q, M\}$ is called

- *distributed* (or space-dependent), if at least one of its system parameters or state variables depend on a space variable.
- *lumped* otherwise.

Example 1.18: Classify examples

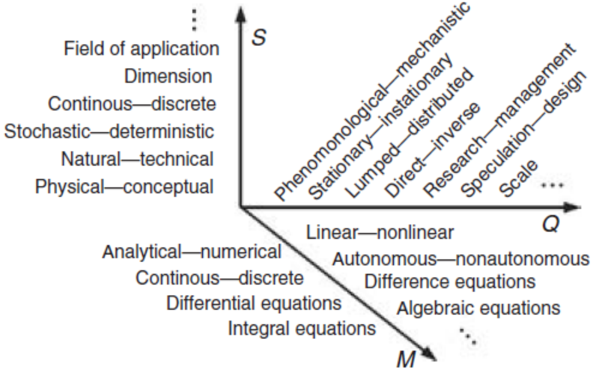
Go through Examples 1.6, 1.8, 1.9, 1.10, 1.11, 1.12 and 1.14. Are they phenomenological or mechanistic? Stationary or instationary? Distributed or lumped?

Types of models

We will consider in particular phenomenological models in Part 2, while we otherwise focus mainly on mechanistic models in Part 3 and 4. Instationary models will play a strong role in Part 3 (ODE models) and distributed models become important in Part 4 (PDE models).

Types of models

Kai Velten considers the following axes for classifying mathematical models:



No need to memorize. But it is in general useful to know which type of system one studies (physical, technical, etc), which type of model is needed to answer the question (mechanistic, instationary, etc) and which type of mathematics (nonlinear, numerical, ODEs, etc) one deals with, as these typically influence the type of methods one has to apply for the simulation step.

Errors of mathematical modeling and simulation

When asking a question, say the *true* answer is x_f.

We take the “detour” via a mathematical model to answer the question, and the mathematical model gives us the solution x_m.

We call the *modeling error* the difference

$$e_m = x_f - x_m.$$

Example 1.19: Example of modeling error

In Example 1.6, the model predicted that our skydiver had a velocity of 98 m/s (=352.8 km/t) after 10 s, while a measurement of a real skydiver would probably not be over 200 km/t (at any time). Hence, we have a substantial modeling error.

Errors of mathematical modeling and simulation

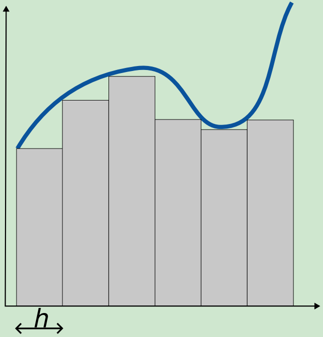
Finding x_m assumes we could perform the mathematical statements of the model exactly. Sometimes we need to use a numerical model to find an approximate solution x_n.

We call the *numerical error* the difference

$$e_n = x_m - x_n.$$

Example 1.20: Example of numerical error

Say the mathematical statements involves calculating an integral where we cannot analytically calculate it. We then rely on approximations via e.g. Riemann sums:



Errors of mathematical modeling and simulation

When using a computer for the numerical model, we are also limited by the computer not being able to represent all numbers. Hence, the computer gave us x_c.

We call the *roundoff error* the difference

$$e_r = x_n - x_c.$$

Example 1.21: Example of roundoff error

If using π in a calculation (e.g. volume/area of cylinder in Example 1.10), the computer cannot use the *actual* value of π, but has an approximate representation of it.

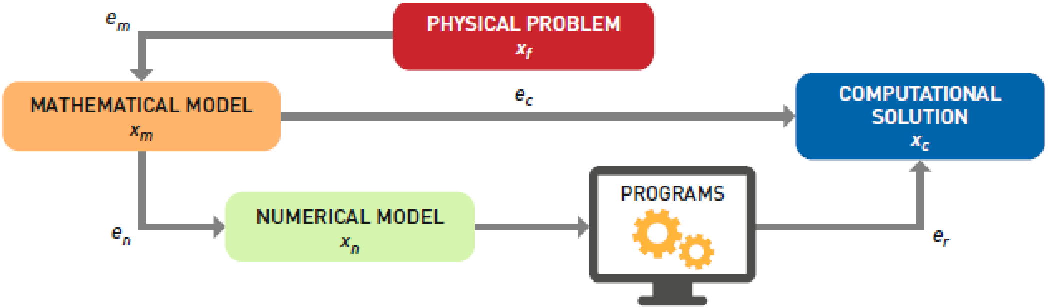
The sum of the numerical and roundoff error is called *computational error*:

$$e_c = e_n + e_r = x_m - x_c.$$

In the end we hope that

$$x_c \simeq x_f.$$

Errors of mathematical modeling and simulation



Errors of mathematical modeling and simulation

Can we control or estimate the errors?

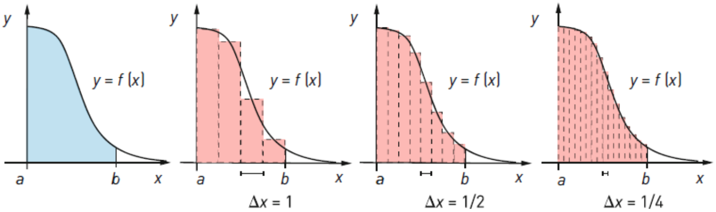
- The roundoff error is usually known and depends on the computer and/or software we use. Typically very small; $\approx 10^{-16}$.
- The numerical error depends on the numerical method that is used. Can usually be quantified using discretization parameter h.
- The modeling error depends on how well our model describes the physical problem. Generally difficult to quantify.

This brings us back to Step 5 from slide 5. We need to *validate* the model.

We validate the (entire) model by investigating whether $x_c \simeq x_f$.

If we used a numerical model, it first makes sense to *verify* the numerical model by investigating whether $x_n \simeq x_m$. This we can do by investigating the error of the numerical method.

Errors of mathematical modeling and simulation



Say the numerical error e_n depends on a discretization parameter h.

Definition 1.22: Convergence of numerical method

A numerical method is said to be *convergent*, or that it converges, when

$$e_n(h) \to 0 \text{ as } h \to 0.$$

The numerical method *converges with order p* (or has convergence of order p) when

$$|e_n(h)| < Ch^p \text{ as } h \to 0$$

for some constant C.

Warning 1.23: Always verify the numerical method!

If you used a numerical method as part of your modeling, always verify the numerical method to make sure you can control the numerical error. Remember that

$$x_f - x_c = e_m + e_c = e_m + e_n + e_r.$$

Hence, to validate the model, it is necessary to verify the numerical model by addressing the numerical error.

To be able to make the final validation of the model, we rely on observational or experimental data of the system we are modeling.

Warning 1.24: Verify or validate?

The terms “verification” and “validation” show up in many contexts and can be difficult to distinguish. For us the main difference can be summarized as:

- Verification answers the question “Did we make the model correctly?”
- Validation answers the question “Did we make the correct model?”

Hence, verification means investigating if each part of the model behaves as intended from a theoretical perspective. Validation means comparing to the real world to check if we get a reasonable answer.