Mathematical Modelling

APM348 Slides*

Bernardo Galvão-Sousa

Exercise 1

1.1 What is modelling?

- A precise description of a system
- A formal summary of knowledge
- A tool that enables prediction
- An abstraction suitable for a particular purpose or question
- Modelling is a scientific method with “hypothesis” in a mathematical form

1.2 Modelling Procedure – DABAR[a]

Step 1. **D**efine the problem

(ask a question)

Step 2. make **A**ssumptions

(select a modelling approach)

Step 3. **B**uild a model

(formulate the model)

Step 4. **A**ssess the model

(solve the model)

Step 5. **R**eport results

(answer the question)

[a] based on the https://m3challenge.siam.org/wp-content/uploads/siam-guidebook-final-press.pdf.

1.3 Course topics:

- Optimization models
- Dynamical models
- Probability models

Optimization Models

Optimization Problem[a]**.** A pig weighting 90 kg gains 3 kg per day and cost 45 cents a day to keep. The market price for pigs is 65 cents/kg, but is falling at 1 cent per day. When should the pig be sold?

[a] Adapted from “Mathematical Modelling” by Meerschaert.

Introduce variables:

- t = time at which the pig is sold (in days)
- w = weight of the pig (in kg)
- m = market price of a pig (in \$/kg)
- C = cost of keeping the pig (in \$)
- R = revenue from selling the pig (in \$)
- P = profit from the sale of the pig (in \$)

2.1 Which of these variables depend on t? Based on the statement, what do we know about their values?

2.2 What is our goal?

2.3 Solve the problem.

2.4 Answer the question: when should the pig be sold and what is the profit?

Parameter Sensitivity.

Parameter sensitivity is a measure of how a model's response is affected by its parameters.
We quantify the **sensitivity** for the model output x and model parameter p by

$$S(x,p) = \frac{\partial x}{\partial p} \cdot \frac{p}{x},$$

which is dimensionless.

Example: If the time to sell or the profit depends strongly on a parameter, then the model is not very useful. If the model said to sell at $t = 1$ if the daily maintenance cost changed to 46 cents, then the recommendation would be very suspect!

2.5 Let $(t^\star, P^\star)$ be the optimal values found before.

What is the sensitivity of P over the parameter c_d = the daily maintenance cost of keeping a pig?

2.6 Is $S(P^\star, c_d)$ positive/negative? What does that mean? Does that make sense?

2.7 What is the sensitivity of P over the parameter m_0 = the initial market price of a pig (in \$/kg)?

2.8 Is $S(P^\star, m_0)$ positive/negative? What does that mean? Does that make sense?

Solutions:

2.1
- $w(t) = 90 + 3t$
- $m(t) = 0.65 - 0.01t$
- $C(t) = 0.45t$
- $R(t) = p(t) \cdot w(t)$
- $P(t) = R(t) - C(t)$

2.2 The goal is to maximize $P(t)$ over $t \geq 0$.

2.3 $P(t) = (90 + 3t)(0.65 - 0.01t) - 0.45t$
$\dfrac{dP}{dt} = 3(0.65 - 0.01t) - 0.01(90 + 3t) - 0.45 = 0$
$t^\star = 10$
$P^\star(10) = 61.50$

2.4 The pig should be sold on day 10, which will give a profit of \$61.50.

2.5 We have $P = (90 + 3t)(0.65 - 0.01t) - c_d t$ so that

$$\begin{aligned} S(P^\star, c_d) &= \frac{\partial P^\star}{\partial c_d} \frac{c_d}{P^\star} \Big|_{c_d = 0.45} \\ &= -t^\star \frac{c_d}{P^\star} \Big|_{c_d = 0.45} = -0.0731707 \end{aligned}$$

This model is insensitive with respect to the maintenance cost! =)

2.6 It is negative, which means that increasing the daily maintenance cost will decrease the profit, which makes sense.

2.7 We get $S(P^\star, m_0) = 1.26829$, so this model is moderately sensitive to the initial price for a pig. =/

2.8 The sensitivity is positive since increasing the initial price of a pig increases the profit also.

Robustness. How do the results depend on the assumptions?

We assumed:

- a linear increase in weight of the pig
- a linear decrease in the price of the pig

What happens if these were nonlinear? The prediction of prices is notoriously uncertain.

Prices are often modelled as stochastic processes (like Brownian motion). This would necessitate a different modelling approach.

In particular, we might then want to maximize the expected (average) profit. But if the variance is very large, then the farmer might prefer a lower expected profit if that means lowering the risk (variance). The farmer might consider maximizing the expected profit with a constraint on the variance of the profit.

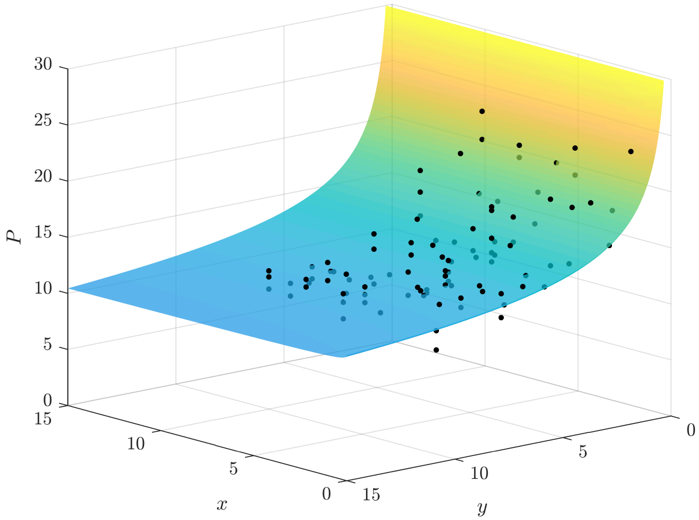
A manufacturer of lawn furniture makes two types of chairs, one with a wood frame and the other with an aluminum frame. The wood frame chair costs \$18 per unit to manufacture and aluminum frame chair costs \$10 per unit to manufacture. The company operates in a market where the number of units that can be sold depends on price. It is estimated that in order to sell x units per day of the wood chair and y units per day of the aluminum chair, the selling price cannot exceed $10+31x^{-0.5}+1.3y^{-0.2}$ dollars per unit for the wood chair and $5+15y^{-0.4}+0.8x^{-0.08}$ dollars per unit for the aluminum chair.

Let us first investigate the selling price model for **one type of** chair.

3.1 As more chairs of both types are sold in the market: $x \to \infty$, what do you expect will happen to their selling price?

3.2 As chairs become scarce: $x \to 0^+$, what happens to the price?

3.3 What family of functions satisfies both these conditions?



Historical prices and fitting surface $p = f(x, y)$.

Exercise 4

A manufacturer of lawn furniture makes two types of chairs, one with a wood frame and the other with an aluminum frame. The wood frame chair costs \$18 per unit to manufacture and aluminum frame chair costs \$10 per unit to manufacture. The company operates in a market where the number of units that can be sold depends on price. It is estimated that in order to sell x units per day of the wood chair and y units per day of the aluminum chair, the selling price cannot exceed $10+31x^{-0.5}+1.3y^{-0.2}$ dollars per unit for the wood chair and $5+15y^{-0.4}+0.8x^{-0.08}$ dollars per unit for the aluminum chair.

4.1 We want to maximize the manufacturer's profit. What is the function to maximize?

4.2 This is a two-dimensional function, so we need to solve the system

$$\frac{\partial f}{\partial x}=0$$
$$\frac{\partial f}{\partial y}=0$$

Write down this system.

4.3 How can we find the solution?

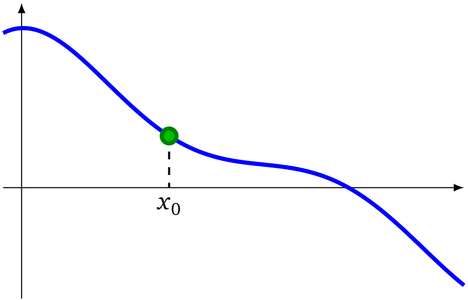
Newton's Method.

This is a method to approximate the solution of the equation

$$f(x) = 0.$$

This is an iterative method, so we start with an initial approximation x_0.
For each successive approximation, take the linear approximation of f at x_i and take x_{i+1} to be the point where the linear approximation is 0.

4.4 From the description above, sketch the point x_1 on the graph on the right when using Newton's method.



4.5 What is the formula for x_1?

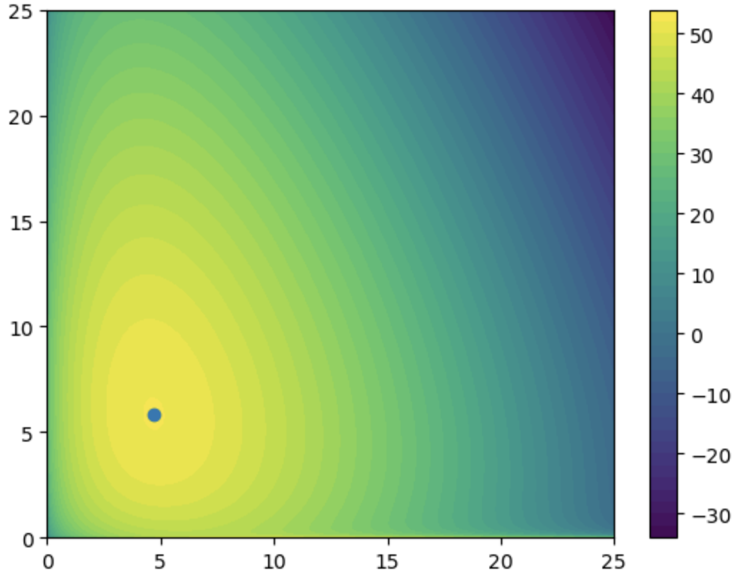
4.6 Leveraging python.

(a) Clone the file `chairs_newton.ipynb` into your Jupyter Notebook

(b) In the file, introduce the partial derivative functions and an initial guess.

(c) Run the script

```
Minimum for  4.689577973016851  wooden chairs and  5.852031046491972  aluminum chairs
Profit =  52.072691798595706
```



Exercise 4

4.6 Leveraging python's minimization tools.

(a) Clone the file `chairs_fmin.ipynb` into your Jupyter Notebook

(b) In the file, introduce the profit function and an initial guess.

(c) Run the script

Exercise 5

A manufacturer of lawn furniture makes two types of chairs, one with a wood frame and the other with an aluminum frame. The wood frame chair costs \$18 per unit to manufacture and aluminum frame chair costs \$10 per unit to manufacture. The company operates in a market where the number of units that can be sold depends on price. It is estimated that in order to sell x units per day of the wood chair and y units per day of the aluminum chair, the selling price cannot exceed $10 + 31x^{-0.5} + 1.3y^{-0.2}$ dollars per unit for the wood chair and $5 + 15y^{-0.4} + 0.8x^{-0.08}$ dollars per unit for the aluminum chair.

Sensitivity. To compute $p^\star$, you can use `chairs_sensitivity.ipynb`.

5.1 How sensitive is the profit to the parameter $c = 10$ (the production cost of the aluminum chair)

$$S(p^\star, c) \approx \frac{p^\star(c+h) - p^\star(c)}{h} \cdot \frac{c}{p^\star(c)}?$$

5.2 How sensitive is the profit to the parameter $b = 0.4$ (the exponent of y in the selling price of the aluminum chair)

$$S(p^\star, b) \approx \frac{p^\star(b+h) - p^\star(b)}{h} \cdot \frac{b}{p^\star(b)}?$$

Note that we are using numerical derivatives, since calculating the partial derivatives analytically is usually impossible.

Constrained Optimization. How do we solve optimization problems with constraints?

Lagrange Multipliers.

We want to minimize (or maximize) a function $f(x)$ with several constraints:

$$g_1(x) = c_1$$
$$\vdots$$
$$g_k(x) = c_k$$

If $x^\star \in \mathbb{R}^N$ is a local optimal of $f(x)$ which satisfies the above constraints, and $\nabla g_1(x^\star), \ldots, \nabla g_k(x^\star)$ are linearly independent, **then**

$$\nabla f(x^\star) = \lambda_1 \nabla g_1(x^\star) + \cdots + \lambda_k \nabla g_k(x^\star), \qquad \text{(LM)}$$

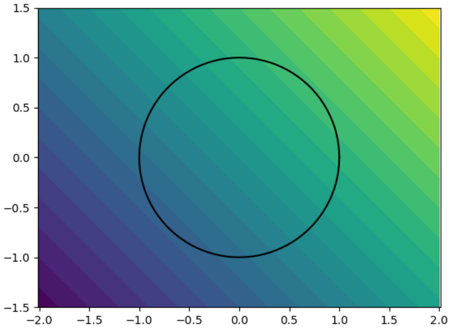
for some scalars $\lambda_1, \ldots, \lambda_k$.

Notes:.

1. This is a necessary, but not sufficient condition.
2. To solve the optimization problem, find candidates x that satisfy it, and then pick the best one.
   - Points for which $\nabla g_1(x), \ldots, \nabla g_k(x)$ are linearly dependent should also be candidates.
3. (LM) $\Leftrightarrow \nabla f(x^\star) \in \operatorname{span}\left\{\nabla g_1(x), \ldots, \nabla g_k(x)\right\}$.
4. The "optimal" values for $\lambda_1, \ldots, \lambda_k$ give important insights on the problem, as we will see – don't ignore them!

Example. Consider the problem:

- Maximize $x + y$ such that $x^2 + y^2 = 1$.



6.1 Use Lagrange Multipliers to find the maximum (and the minimum).

6.2 If the constraint was $x^2 + y^2 = c$, then what is:

(a) the maximizer point $(x^\star, y^\star)$?

(b) the Lagrange multiplier $\lambda^\star$?

(c) the maximum $f(x^\star, y^\star)$?

6.3 Compare $\lambda^\star$ with $\dfrac{\partial f(x^\star, y^\star)}{\partial c}$.

6.4 Based on this relation, give an interpretation for the Lagrange Multiplier.

6.1

$$\nabla f = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{and} \quad \nabla g = \begin{bmatrix} 2x \\ 2y \end{bmatrix}$$

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \lambda \begin{bmatrix} 2x \\ 2y \end{bmatrix} \iff \begin{cases} 1 & = 2\lambda x \\ 1 & = 2\lambda y \\ 1 & = x^2 + y^2 \end{cases}$$

$$1 = \frac{1}{2\lambda^2} \iff \lambda = \pm\frac{1}{\sqrt{2}}$$

$$x = y = \pm\frac{1}{\sqrt{2}}$$

6.2 $x^\star = y^\star = \frac{\sqrt{c}}{\sqrt{2}}$ and $\lambda^\star = \frac{1}{\sqrt{2c}}$

$$\max = x^\star + y^\star = \sqrt{2c}$$

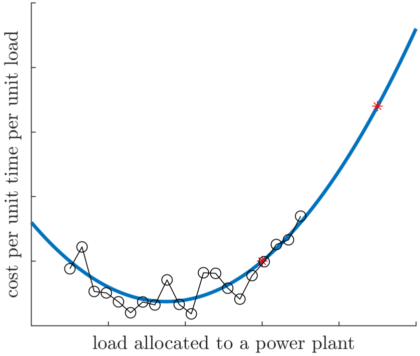
6.3 $\dfrac{\partial f(x^\star, y^\star)}{\partial c} = \dfrac{\sqrt{2}}{2\sqrt{c}} = \lambda^\star$

6.4 This means that if the constraint increased from 1 to $1+\Delta = 1.1$, then we would expect the maximum to increase by approximately $\Delta\lambda^\star = \frac{\Delta}{\sqrt{2}} \approx 0.07$.

Indeed, $\Delta f = \sqrt{2.2} - \sqrt{2} \approx 0.069$.

Define the problem.

> The production side of the electrical power grid[a] consists of hundreds or thousands of power plants that vary in fuel sources (coal, nuclear, hydroelectric, solar, wind, stored energy in the batteries of electric vehicles, etc.) and characteristics (age, efficiency, automated, etc.).
> How can the power consumption load be allocated to these plants to minimize cost?
>
> [a] This example is based on Huijuan Li in 'Lagrange Multipliers and their Applications'.



Make Assumptions.

- Each power plant is summarized by a cost curve which tells how much a given load costs. Generally, the cost per unit time per unit load of operating a power plant is a concave function of load as in the figure below: small and large loads are expensive.
- For simplicity, we will approximate these quadratics by a linear function with one parameter: the cost per unit time per unit load is $c(x) = ax + 1$, so the cost rate function has the form $f(x) = (ax+1)x = ax^2 + x$.

- N = number of power plants
- x_i = load assigned to power plant i (in MW)
- X = total load (in MW)(In Toronto the average total load is 2500 MW.).
- C = cost rate of power generation (in \$/h)
- $f_i(x_i)$ = cost rate function for power plant i (in \$/h)

Build a model.

7.1 Find an equation relating X and x_i.

7.2 Find a formula for C.

7.3 Formulate the problem we want to solve.

Assess the model.

We are going to assume the following:

- Three power plants identified with the parameters:
  - $a_1 = 0.0625$
  - $a_2 = 0.0125$
  - $a_3 = 0.0250$
- The total load is 925 MW

7.4 Solve the problem.

Report the results.

7.5 What is the interpretation of $\lambda^\star$ the "optimal" Lagrange multiplier?

7.6 What is the sensitivity of the cost with respect to the parameters a_i and X? What does that mean about the model?

7.3 Objective: $\min \sum_{i=1}^{3} a_i x_i^2 + x_i$

Constraint: $\sum_{i=1}^{3} x_i = X$

7.4 Define:

$$C(\vec{x}) = \sum_{i=1}^{3} a_i x_i^2 + x_i$$

$$g(\vec{x}) = \sum_{i=1}^{3} x_i = X$$

So we have

$$\nabla C(\vec{x}) = \begin{bmatrix} 2a_1 x_1 + 1 \\ 2a_2 x_2 + 1 \\ 2a_3 x_3 + 1 \end{bmatrix} = \lambda \nabla g(\vec{x}) = \lambda \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Which can be written as

$$\begin{bmatrix} 2a_1 & 0 & 0 & -1 \\ 0 & 2a_2 & 0 & -1 \\ 0 & 0 & 2a_3 & -1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \lambda \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ -1 \\ X \end{bmatrix}$$

And we get the unique solution:

- $x_1 = 112$ MW
- $x_2 = 560$ MW
- $x_3 = 280$ MW
- $\lambda = \$15$ /h/MW (shadow cost)

We used: `power-plants.ipynb`

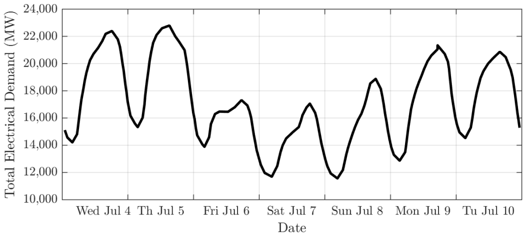
7.5 If we reduce the total load (X) by 1 MW, it would approximately reduce the total cost of operating the three power plants by \$15/h.

So the operator of the power plants should be willing to pay consumers who pump electricity back to the grid up to \$15/h for each megawatt.

7.6
- $S(C, X) \approx 1.875$
- $S(C, a_1) \approx 0.000015$
- $S(C, a_2) \approx 0.00017$
- $S(C, a_3) \approx 0.00007$

Robustness.

8.1 The parameter X varies significantly (regularly by over 50% in a day), so understanding it is very important.



It is crucial to understand how the optimal cost and loads change with X.

8.2 Is the quadratic model for f_i good? You can try different functions.

8.3 Should there be other constraints on x_i? We only imposed $x_i > 0$, but we probably should impose upper bounds too.

8.4 What about transportation costs? There can be losses of up to 20% on high-tension transmission lines.

8.5 We have a static model, where the power plants operate always at the same load. We might want to consider a dynamic optimization model.

Exercise 9

Linear Programming[a]**.** A family farm has 1250 hectares[b] of land for planting. Possible crops that they could plant are corn, wheat, and oats. There are 400 hectare-m (a volume) of water available for irrigation and 600 hours of labour per week available. The requirements and expected yields are shown below.

	corn	wheat	oats
irrigation (ha-m / ha)	1.0	0.3	0.5
labour (person-h / week / ha)	1.6	0.4	0.6
yield ($/ha)	1400	420	700

We want to maximize the total yield.

[a] based on a problem from Meerschaert's 'Mathematical Modeling'.
[b] 1 hectare = 1 ha = $10\,000$ m^2.

Introduce the following variables:

- x_i = hectares planted of $i = 1$ corn, $i = 2$ wheat, $i = 3$ oats
- w = the total irrigation used in ha-m
- ℓ = the total labour used in person-h / week
- a = the total area planted in hectares
- y = the total yield in $

9.1 Find expressions for w, ℓ, a, y

9.2 What are the constraints on the variables defined?

9.3 Formulate the optimization problem we want to solve in standard linear programming form:

$$\begin{array}{ll} \text{Objective:} & \max\ \vec{c}^T\vec{x} \\ \text{Constraints:} & \begin{array}{c} A\vec{x} \le \vec{b} \\ \vec{x} \ge \vec{0} \end{array} \end{array}$$

9.4 Use `farm-linearprog.ipynb` to find the solution.

Exercise 9

9.1
- $w = 1x_1 + 0.3x_2 + 0.5x_3$
- $\ell = 1.6x_1 + 0.4x_2 + 0.6x_3$
- $a = \sum_{i=1}^{3} x_i$
- $y = 1400x_1 + 420x_2 + 700x_3$

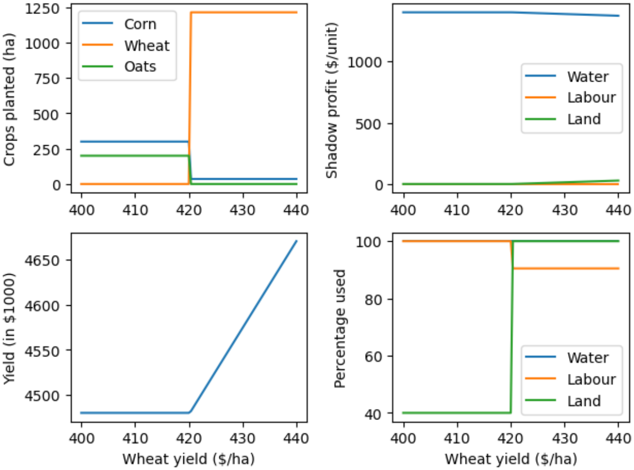
9.2
- $x_i \geq 0$
- $w \leq 400$
- $\ell \leq 600$
- $a \leq 1250$

9.3 Objective: $\max \begin{bmatrix} 1400 & 420 & 750 \end{bmatrix} \vec{x}$

Constraints:
$$\begin{bmatrix} 1 & 0.3 & 0.5 \\ 1.6 & 0.4 & 0.6 \\ 1 & 1 & 1 \end{bmatrix} \vec{x} \leq \begin{bmatrix} 400 \\ 600 \\ 1250 \end{bmatrix}$$
$$\vec{x} \geq \vec{0}$$

Exercise 9

We ran the same model with the Wheat Yield ranging from \$400/ha to \$440/ha and obtained the following graphs.



9.5 Interpret the results and the shadow profit (— shadow cost).

Modified farming problem. We modify the original optimal farming problem to include the notion of plots. The 1250 hectares farm is broken down into 5 plots of 240 hectares each and one 50 hectare plot. For convenience, the farmers want to plot only one crop on each plot. As before, 400 ha-m of water and 600 hours of labour are available. The requirements and expected yields are shown below.

	corn	wheat	oats
irrigation (ha-m / ha)	1.0	0.3	0.5
labour (person-h / week / ha)	1.6	0.4	0.6
yield ($/ha)	1400	420	700

We want to maximize the total yield.

Introduce the variables:

- x_1, x_2, x_3 are the number of large plots of corn, wheat, and oats respectively;
- x_4, x_5, x_6 are the number of small plots of corn, wheat, and oats respectively.

10.1 Set up and solve the problem.

10.2 Interpret the results.

Ice Cream[a].
Suppose a manufacturing company receives an order for B units to be delivered at time T, e.g. Sobeys has placed an order for $B = 100$ pallets of Chapman's vanilla ice-cream for a promotion starting in $T = 10$ days.
Chapman's Ice Cream must decide when to produce their tasty product. They don't want to produce it early since they will have to pay to keep it frozen until the order is due. They also do not want to produce it the day before it is due since running the production line fast might have a large cost.

[a] Based on an example from Kamien and Schwartz's 'Dynamic Optimization'

Let $x(t)$ be the inventory at time t and suppose that $x(0) = 0$ and to fill the order we need $x(T) = B$ (boundary conditions).

11.1 Let us divide the time interval $[0, T]$ into N "chunks". What is the length Δt of each?

11.2 Let Δx_n be the number of units produced during the n^{th} time interval. Find a formula relating Δx_n with $x(t)$. Find an equation relating Δx_n with B.

11.3 We need to consider the cost of storing the produced units in inventory: assume that each unit has a cost of c_2 per unit time. What is the total inventory cost?

11.4 We want to model the fact that running machines faster is more costly. What is a model for the cost of producing Δx_n units during a time interval of length Δt that quantifies this?

11.5 What is the total production cost?

11.6 What is the total cost?

11.7 What are the constraints for the variables?

11.8 Approximate the solution.

Exercise 11

11.1 Let us break the time interval $[0, T]$ into $\Delta t = T/N$ "chunks" and consider $t_n = n\Delta t$. We need to decide how many units Δx_n to produce at each time interval.

11.2 We then have:

- $x(t_{n+1})) = x(t_n) + \Delta x_n$
- $\Delta x_1 + \cdots + \Delta x_N = B$

11.3 We need to consider the cost of storing the produced units in inventory: assume that each unit has a cost of c_2 per unit time:

- Inventory Cost $= \displaystyle\sum_{n=1}^{N} \Delta x_n (T - t_n) c_2$

11.4 If the production cost was: $\displaystyle\sum_{n=1}^{N} c\Delta x_n$, then $c =$ the cost of producing 1 unit in Δt time. If this is constant, then there is no penalty in running the machines faster, so we need to consider c that is not constant and depends on Δx_n: we make the modelling assumption $c = c_1 \frac{\Delta x_n}{\Delta t}$, so that c is proportional to the rate of production. We get

- Production Cost $= \displaystyle\sum_{n=1}^{N} \frac{\Delta x_n^2}{\Delta t} c_1$

11.5 So the total cost is

- Total Cost $= \displaystyle\sum_{n=1}^{N} \left[\Delta x_n^2 c_1 + \Delta x_n (N - n) c_2 \right]$

11.6 The constraints are

- $\Delta x_1 + \cdots + \Delta x_N = B$
- $\Delta x_n \geq 0$

11.7 The solution is here: `IceCream.ipynb`

Exercise 12

In the previous problem, instead of modelling it using **discrete time**, we can model it using **continuous time**.

Then, we have the following:

- $\frac{dx}{dt}(t) =$ units produced per unit time (at time t)
- Inventory cost $= \displaystyle\int_0^T c_2 \frac{dx}{dt}(t)(T-t)\,dt = \int_0^T c_2 x(t) dt$ (why?)
- Production cost $= \displaystyle\int_0^T c_1 \left(\frac{dx}{dt}\right)^2 dt$ (why?)

We can formulate the problem as

Objective: $\displaystyle\min \int_0^T c_1(x'(t))^2 + c_2 x(t)\,dt$

Constraints:
$$x(0) = 0 \text{ and } x(T) = B$$
$$x'(t) \geq 0$$

The goal here is to find a function $x(t)$. This is a problem in **Calculus of Variations**.

Euler-Lagrange Equation.
We want to find a function $x : [t_0, t_1] \to \mathbb{R}$ that minimizes the functional:

$$\min \int_{t_0}^{t_1} F(t, x(t), x'(t))\, dt$$

and $x(t_0) = x_0$ and $x(t_1) = x_1$.

When we want to find a minimizer of a function, *we set the derivative to zero.*

13.1 The definition of derivative for a real function is

$$f'(x) = \lim_{\varepsilon\to 0} \frac{f(x+\varepsilon) - f(x)}{\varepsilon}$$

We only have one direction for ε, so this limit suffices. For a function of multiple variables, we introduced the notion of partial derivative:

$$\frac{\partial f}{\partial x_i}(\vec{x}) = \lim_{\varepsilon\to 0} \frac{f(\vec{x} + \varepsilon\vec{e}_i) - f(\vec{x})}{\varepsilon}$$

Our case is similar, but instead of having vectors as inputs, our inputs are functions $x(t)$, so our definition must be adapted to:

- Let $y(t) = x(t) + \varepsilon v(t)$

What are conditions on $v(t)$ that guarantee that $y(t)$ is an admissible function for the problem formulated in the blue box above?

13.2 Let $g(\varepsilon) = \displaystyle\int_{t_0}^{t_1} F(t, y(t), y'(t))\, dt$. Expand the formula for $g(\varepsilon)$.

13.3 Expand $g'(0)$.

13.4 Set $g'(0) = 0$ and solve.

Hint: If $\int_a^b f(t)v(t)\, dt = 0$ for every function $v(t)$ satisfying $v(a) = v(b) = 0$, then $f(t) = 0$ for all $t \in (a, b)$.

Euler-Lagrange Equation.
The minimizer $x^\star(t)$ of the functional

$$\min \int_{t_0}^{t_1} F(t, x(t), x'(t))\, dt$$

with $x(t_0) = x_0$ and $x(t_1) = x_1$ satisfies the **Euler-Lagrange Equation**:

$$\frac{\partial F}{\partial x}(t, x^\star, x^{\star\prime}) = \frac{d}{dt}\frac{\partial F}{\partial x'}(t, x^\star, x^{\star\prime}).$$

We will look back to **Exercise 12**.

14.1 Use the Euler-Lagrange Equation to obtain a Differential equation for $x(t)$.

14.2 Solve the differential equation with the boundary conditions.

14.3 We required $x'(t) \geq 0$. Does this solution satisfy this condition?

14.4 To get a solution that satisfies $x' \geq 0$, we need to consider a solution that doesn't produce any units for a while:

$$x(t) = \begin{cases} 0 & \text{if } t < t_1 \\ z(t) & \text{if } t_1 \leq t \leq T \end{cases}$$

What is t_1 and what is the function $z(t)$?

14.5 If we add a constraint $x'(t) \leq M$, how would that modify the solution? What does this restriction mean in the ice-cream context?

14.1

$$\frac{\partial F}{\partial x} = c_2$$
$$\frac{\partial F}{\partial x'} = c_1 2x'(t)$$
$$\frac{d}{dt}\frac{\partial F}{\partial x'} = 2c_1 x''(t)$$

So the Euler-Lagrange equation yields $x''(t) = \frac{c_2}{2c_1}$.

14.2 The general solution of the ODE is: $x(t) = \frac{c_2}{4c_1}t^2 + v_0 t + x_0$

Using the boundary conditions we get:

$$x(t) = \frac{c_2}{4c_1}t^2 + \frac{4c_1 B - c_2 T^2}{4c_1 T}t$$

14.3 If $B < \frac{c_2 T^2}{4c_1}$, then x' can be negative at the beginning:

$$x'(t) \leq 0 \Leftrightarrow \frac{c_2}{2c_1}t + \frac{4c_1 B - c_2 T^2}{4c_1 T} \leq 0$$
$$\Leftrightarrow t \leq \frac{c_2 T^2 - 4c_1 B}{c_2 T}$$

This only happens for small values of B. Intuitively, this means that since the order is small, the producer would be better off by selling more of their product to save on inventory (inventory cost becomes negative) and produce the required order later.

14.4 The solution is is decreasing when $c_2T^2 - 4c_1B > 0$, so to make sure that this doesn't happen for the new solution, we choose t_1 such that $c_2(T-t_1)^2 - 4c_1B = 0$:

$$t_1 = T - \sqrt{\frac{4c_1B}{c_2}}$$

The function $z(t)$ is the optimal function $x(t)$ just translated by t_1 and with $T \to T - t_1$:

$$z(t) = \frac{c_2}{4c_1}(t-t_1)^2 + \frac{4c_1B - c_2(T-t_1)^2}{T-t_1}(t-t_1)$$

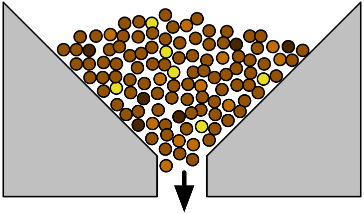
`https://www.desmos.com/calculator/ny2frmc2ov`

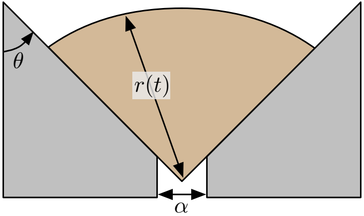
14.5 If B is not too large: $B \leq MT - \frac{c_2}{4c_1}T^2$, then the original solution holds.

If B is too large, then we have too many units to produce in the time provided, so we would need to produce as many as we could ($x'(t) = M$) at the end to be able to complete the order. Before that time, we could produce at the optimal rate.

`https://www.desmos.com/calculator/2rfh1w2a7a`

Dynamical Models





The following ordinary differential equation models a crowd leaving a stadium through an exit

$$2\theta r \frac{dr}{dt} = -k\alpha\sqrt{r}$$

based on the premise

(TL) Torricelli's Law: The area of the region occupied by the crowd decreases proportionally to the width of the exit times the square root of its radius.

16.1 How is the premise expressed in the differential equation?

16.2 Sketch a slope field for this model

https://www.desmos.com/calculator/1xb4g6cuiz

and use it to study how the time it would take to evacuate that section depends on the parameters.

16.3 Using Euler's method, estimate how long it would take to evacuate a stadium with $\alpha = k = 1$, $\theta = \frac{\pi}{5}$ and $r(0) = 2$.

Exercise 17

Ladd Peebles Stadium

According to the paper "`A study of stadium exit design on evacuation performance`" studying the Ladd Peebles stadium:

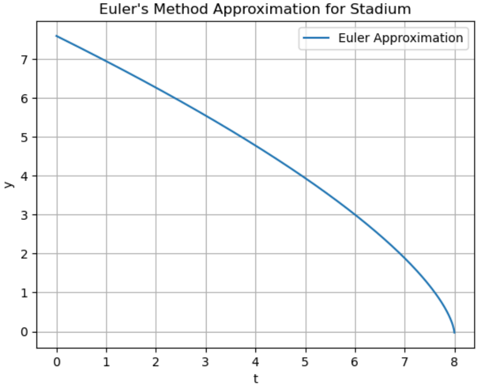
- The average person occupies 0.15m^2.
- The stadium fits 1200 people in one section.
- The exits are 1.5m wide.

17.1 According to an experiment in the paper, it took 8 minutes to evacuate the stadium. Use this to estimate k for Ladd Peebles.

17.2 In the same paper, "for safety, the maximum flow through an exit is 109 people per meter-width per minute." Does Ladd Peebles satisfy this safety concern?

Solution:

- $\theta r^2(0) = 1200 \cdot (0.15) \quad \Rightarrow \quad r(0) \approx 7.6m$
- $\theta = \pi$
- $\alpha = 1.5$
- To get everyone out in 8 minutes $\Rightarrow k = 7.33$ (time units are minutes)
- $p(t) = A(r(t))/(0.15 \cdot 1.5) =$ people per meter-width
- $p(t) = 2\theta \frac{1}{2} r^2(t)/(0.15 \cdot 1.5) = \frac{\theta}{0.225} r^2(t)$
- $p'(t) = \frac{1}{0.225} \underbrace{2\theta r \frac{dr}{dt}}_{-k\alpha\sqrt{r}} = -\frac{k\alpha}{0.225}\sqrt{r(t)} = -\frac{152}{3}\sqrt{r(t)}$
- Max at $t = 0$ when $|p'(t)| \approx 139.678$
- The solution is here: `Stadium-Euler.ipynb`



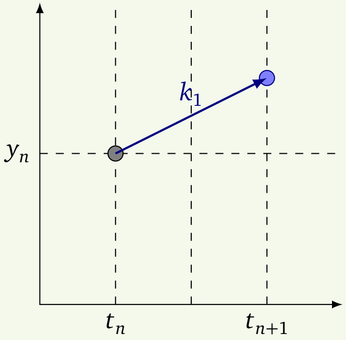

```
Max sqrt(y) is  2.756809750418044
```

Numerical Methods for:

$$y' = f(t, y)$$

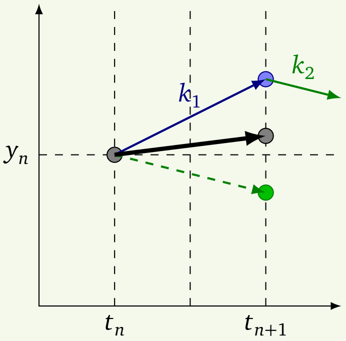
Euler Method.

$$y_{n+1} = y_n + hk_1$$
$$k_1 = f(t_n, y_n)$$



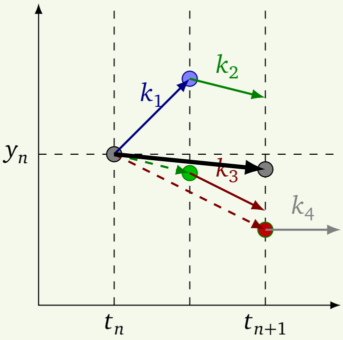
Heun Method (Improved Euler).

$$y_{n+1} = y_n + h\frac{k_1 + k_2}{2}$$
$$k_1 = f(t_n, y_n)$$
$$k_2 = f(t_n + h, y_n + hk_1)$$



Runge-Kutta Method (4^{th} order).

$$y_{n+1} = y_n + h\frac{k_1 + 2k_2 + 2k_3 + k_4}{6}$$
$$k_1 = f(t_n, y_n)$$
$$k_2 = f(t_n + \frac{h}{2}, y_n + \frac{h}{2}k_1)$$
$$k_3 = f(t_n + \frac{h}{2}, y_n + \frac{h}{2}k_2)$$
$$k_4 = f(t_n + h, y_n + hk_3)$$



Desmos with all these three methods: `https://www.desmos.com/calculator/haolaltd9s`

Consider the ODE $\frac{dy}{dx} = 2x\sin(x^2)$.

18.1 Recall the meaning of the line segments in the slope field for this ODE.

18.2 Consider the solution satisfying $y(0) = 0$. With a step $h = 0.1$, find the largest interval that the approximations stay within 0.1 distance of the exact solution.

Exercise 18

The exact solution is

$$y = 1 - \cos(x^2).$$

And by observing it on Desmos:

`https://www.desmos.com/calculator/qflikqjufs`

We conclude that

- Euler: $x < 1.2$
- Heun: $x < 5.6$
- Runge-Kutta: all x ?

Dimensional Analysis

Seven Fundamental Dimensions.

There are seven fundamental dimensions:

Dimension	Symbol	SI Unit	
length	L	metre	m
mass	M	kilogram	kg
time	T	second	s
electric current	I	ampere	A
temperature	Θ	kelvin	K
amount	N	mole	mol
light intensity	J	candela	cd

Note: Sometimes, we use charge Q (SI Unit coulomb C) as a fundamental dimension instead of current.

19.1 When can we add/subtract quantities? With different dimensions? With the same dimensions?

19.2 When can we equate quantities? With different dimensions? With the same dimensions?

19.3 When can we multiply/divide quantities? With different dimensions? With the same dimensions?

19.4 It is convenient to define some functions as a power series (e.g. $e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \cdots$). What condition on the dimension of x is required to be able to do this?

19.5 What are the dimensions of a derivative $\frac{dy}{dx}$? What are the dimensions of an integral $\int y\,dx$?

Modelling: Relationship between the variables in a model must be dimensionally consistent.

Non-Dimensionalization. Consider the model for a mass undergoing radioactive decay:

$$\frac{dm}{dt} = -km$$

with $m(0) = m_0$.

20.1 What are the units of k? What are the units of $t_c = \frac{1}{k}$?

20.2 Introduce new variables: $\tau = \frac{t}{t_c}$ and $\overline{m}(\tau) = \frac{m(t)}{m_0}$. What is the ODE satisfied by $\overline{m}(\tau)$? What are its units? What are the parameters for this equation?

20.1 The units of m are mass M, so the units of $\frac{dn}{dt}$ are $\frac{M}{T}$.

This means that the units of k must be $\frac{1}{T}$, so that km matches the units on the other side of the equation.

This implies that t_c has the units of time T.

20.2 $\dfrac{d\overline{m}}{d\tau} = \dfrac{1}{m_0}\dfrac{dm}{d\tau} = \dfrac{1}{m_0}\dfrac{dm}{dt}\dfrac{dt}{d\tau} = \dfrac{t_c}{m_0}\dfrac{dm}{dt}$

So we get

$$\frac{d\overline{m}}{d\tau} = \frac{t_c}{m_0}\frac{dm}{dt} = -\frac{t_c}{m_0}km(\tau) = -\frac{1}{m_0}m(\tau) = -\overline{m}$$

and $\overline{m}(0) = 1$.

Spruce Budworm Outbreak. Consider the model for spruce budworm outbreak in Eastern Canada.[a]

$$\frac{dN}{dt} = RN\left(1 - \frac{N}{K}\right) - \frac{BN^2}{A^2 + N^2}.$$

The first term accounts for resource-limited population growth within a tree and the second term accounts for the predation of the budworms by birds.

[a] See "Nonlinear Dynamics and Chaos" by Strogatz.

21.1 What are the units of N, A, B, K?

21.2 To "non-dimensionalize" this ODE, what variable would you consider instead of N? What ODE is satisfied by your new variable? How many parameters do you have now?

21.1

- $[N] =$ budworm population (N)
- $[K] =$ carrying capacity of budworm population (N)
- $[R] = \frac{1}{T}$
- $[A] = N$
- $[B] = \frac{N}{T}$

21.2 Consider the new variables[a]:

- $x = N/A$ the non-dimensional budworm population
- $\tau = \frac{Bt}{A}$ the non-dimensional time
- $r = \frac{RA}{B}$ the non-dimensional growth rate
- $k = \frac{K}{A}$ the non-dimensional carrying capacity

$$\frac{dx}{d\tau} = \frac{1}{A}\frac{dN}{dt}\frac{dt}{d\tau} = \frac{1}{B}RN\left(1 - \frac{N}{K}\right) - \frac{N^2}{A^2 + N^2}$$
$$= \frac{1}{B}ARx\left(1 - A\frac{x}{K}\right) - \frac{x^2}{(1+x^2)}$$
$$= rx\left(1 - \frac{x}{k}\right) - \frac{x^2}{(1+x^2)}$$

OR consider the new variables:

- $x = N/K$ non-dimensional budworm population (fraction of its carrying capacity)
- $b = B/K$ with units $1/(\text{amount}^2 \times$ time
- $a = A/K$ non-dimensional

$$\frac{dx}{d\tau} = \frac{1}{K}\frac{dN}{dt}\frac{dt}{d\tau} = Rx(1-x) - \frac{1}{K}\frac{BN^2}{A^2 + N^2}$$
$$= Rx(1-x) - \frac{bx^2}{a + x^2}$$

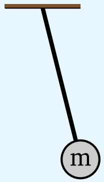
[a] This is not the only way to do this.

Dimensional Matrix. The dimensional matrix $\mathcal{D}$ is a matrix where its (i, j) entry gives the power of the i^{th} dimension of the j^{th} variable.

Buckingham Pi Theorem. Any physical relation involving N dimensional variables can be written in terms of a complete set of $N - r$ independent dimensionless variables, where r is the rank of the dimensional matrix $\mathcal{D}$.
The notational convention for the Buckingham Pi Theorem is that the "pi's", $\Pi_1, \ldots, \Pi_{N-r}$ represent dimensionless variables and a relation between them is given by $F(\Pi_1, \ldots, \Pi_{N-r}) = 0$.

Consider a pendulum. We make assumptions:

- The pivot is frictionless
- The rod is massless
- Air resistance is neglected
- The ceiling is infinitely rigid
- $\ldots$



22.1 What are the units of the following variables of interest?

(a) Period of the swing $[P] =$

(b) Pendulum mass $[m] =$

(c) Pendulum rod length $[\ell] =$

(d) Gravitational acceleration $[g] =$

(e) Amplitude of the swing $[\Theta] =$

Exercise 22

22.2 Let us create the dimensional matrix:

- One column for each variable of interest (remember the order used for later)
- One row for each dimension
- Each term contains the power of the corresponding dimension for the corresponding variable

$$\begin{array}{ccc}
 & \begin{array}{ccccc}[P] & [m] & [\ell] & [g] & [\Theta] \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow\end{array} & \\
\mathcal{D} = & \begin{bmatrix} \quad & \quad & \quad & \quad & \quad \\ & & & & \\ & & & & \end{bmatrix} & \begin{array}{l}\leftarrow M\\ \leftarrow L \\ \leftarrow T\end{array}
\end{array}$$

22.3 What is the rank of this matrix?

22.4 What is the dimension of the null space?

22.5 Find a basis for the null space.

For each vector of the null space basis,

$$\begin{bmatrix} 2 \\ 0 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Buckingham Pi Theorem states that these correspond to non-dimensional variables Π_1 and Π_2:

$$\Pi_1 = \frac{P^2 g}{\ell} \quad \text{and} \quad \Pi_2 = \Theta$$

and that there is a relation between them:

$$F(\Pi_1, \Pi_2) = 0 \quad \text{or} \quad \Pi_1 = f(\Pi_2) \quad \Leftrightarrow \quad \frac{P^2 g}{\ell} = f(\Theta)$$

which implies that

$$P = \sqrt{\frac{\ell}{g}} \cdot \overline{f}(\Theta),$$

or in other words, the fact that the *period of the pendulum is proportional to the square root of its length* is a consequence of a pure dimensional analysis of the variables in the problem.

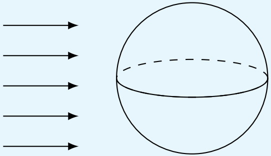
22.6 Recall the ODE for the pendulum: $\frac{d^2\theta}{dt^2} = -\frac{g}{\ell}\sin(\theta)$. Linearize[a] it near the equilibrium $\theta = 0$.

22.7 Solve the linearized pendulum ODE, and compare the period of the linearized model to the nonlinear one.

[a] If you are not comfortable with linearization of an ODE, check exercise 61 on `https://raw.githubusercontent.com/siefkenj/IBLODEs/main/dist/odes.pdf`.

Exercise 23

Consider the flow past a sphere.



You don't need to know much about fluid dynamics to be able to deduce some properties of the flow.

The sphere is in a fluid (water) and we measure the force necessary to keep the sphere from moving downstream.
We want to understand how the drag force depends on the upstream velocity.

23.1 What are the units of the variables of interest[a]?

(a) drag force $[F] =$

(b) upstream velocity $[v] =$

(c) fluid density $[\rho] =$

(d) sphere diameter $[D] =$

(e) fluid viscosity[b] $[\mu] =$

23.2 Create a dimension matrix $\mathcal{D}$.

23.3 What is its rank? What is the dimension of its null space? Find a basis for its null space.

23.4 What are the non-dimensional variables Π's from Buckingham Pi Theorem?

23.5 What relations do you obtain?

[a]This choice is part of the modelling process.

[b]Fluid viscosity is the sphere's resistance to deformation by shear stress. To help with the units, the formula for the Force from viscosity is $F = \mu \cdot A \cdot u / y$, where A is area, u is velocity and y is position.

Solution:

$$\mathcal{D} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & -3 & 1 & -1 \\ -2 & -1 & 0 & 0 & -1 \end{bmatrix}$$

for rows M, L, T.

Its rank is 3, so there are 2 independent null vectors:

$$\begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \\ -1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 \\ -2 \\ -1 \\ -2 \\ 0 \end{bmatrix}$$

corresponding to

$$\Pi_1 = \frac{\rho v D}{\mu} \quad \text{and} \quad \Pi_2 = \frac{F}{\frac{1}{2}\rho v^2 D^2}$$

- $\Pi_1 =$ Reynolds number (Re) which determines the relation between inertia and viscous forces in a fluid flow.
- $\Pi_2 =$ is the drag coefficient (C_d)

So dimensional analysis reveals:

$$\Pi_2 = f(\Pi_1)$$

which means that the drag coefficient depends on the fluid's Reynolds number.

Could have also obtained

$$\begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \\ -1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ -2 \end{bmatrix}$$

which gives a different Π_2 and a different relation.

Exercise 23

Using python to find the null space gives yet another set of different Π_1 and Π_2.

```
import numpy as np
from numpy.linalg import matrix_rank
from sympy import Matrix, nsimplify

D = np.array([[1,0,1,0,1],[1,1,-3,1,-1],[-2,-1,0,0,-1]])
Ds = Matrix([[1,0,1,0,1],[1,1,-3,1,-1],[-2,-1,0,0,-1]])

print(D)

print("\nRank(D)=",matrix_rank(D))

print("\nNull Space Basis for D is \n",-2*nsimplify(Ds, rational=True).nullspace(
```

```
[[ 1  0  1  0  1]
 [ 1  1 -3  1 -1]
 [-2 -1  0  0 -1]]

Rank(D)= 3

Null Space Basis for D is
 Matrix([[1], [-2], [-1], [-2], [0]])
 Matrix([[1], [0], [1], [0], [-2]])
```

Exercise 24

24.1 Use Buckingham Pi Theorem on Exercise 20 about radioactive decay.

24.2 Use Buckingham Pi Theorem on Exercise 21 about the budworm population.

Dog Shampoo. Scientists are testing the effect of different dog shampoos. Let

- F = number of fleas (in millions)
- D = number of dogs (in thousands)
- a = effect of different dog shampoos

and consider the model:

$$F' = -(1+a)F + D - 2$$
$$D' = -2F + (1-a)D + 1$$

which is based on the following premises:

($P1_F$) Ignoring all else, the number of parasites decays in proportion to its population (with constant $1+a$).

($P2_F$) Ignoring all else, parasite numbers grow in proportion to the number of hosts (with constant 1).

($P1_D$) Ignoring all else, hosts numbers grow in proportion to their current number (with constant $1-a$).

($P2_D$) Ignoring all else, host numbers decrease in proportion to the number of parasites (with constant 2).

($P1_C$) Anti-flea collars remove 2 million fleas per year.

($P2_C$) Constant dog breeding adds 1 thousand dogs per year.

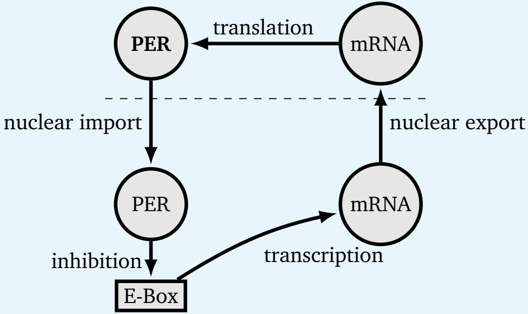
25.1 How are the premises expressed in the differential equations?

25.2 Find the equilibrium solutions for each value of $-1 \le a \le 1$.

25.3 Use `fleas_dogs.ipynb` and eigenvalues to check the stability[a] of the equilibrium points for different values of $-1 \le a \le 1$.

[a] If you are not comfortable with studying the stability of the equilibrium solutions of a system of ODEs, then check exercises 32–61 of the `MAT244 in-class exercises`. You can also check sections 2.4 and 2.5 of the textbook ‘`Diffy Qs`’ `by Jiri Lebl`.

Mammalian Circadian Clock.



When the enhancer-box (E-Box) on the DNA is active, messenger RNA (mRNA) is produced. The mRNA is exported from the nucleus where it is translated into PER protein. The protein is imported into the nucleus where it inhibits the E-Box. It is the cytosolic concentration of PER (highlighted) that indicates the time of day.

We get the model:

- x_1 = enhancer box on the DNA (E-box)
- x_2, x_3 = mRNA inside/outside the nucleus
- x_4, x_5 = PER outside/inside the nucleus

We get:

$$x_1' = -x_1 + e^{-\alpha x_5} \qquad x_4' = -x_4 + x_3$$
$$x_2' = -x_2 + x_1 \qquad x_5' = -x_5 + x_4$$
$$x_3' = -x_3 + x_2$$

where the exponential term represents the fact that the PER protein inhibits the E-box with "strength" α.

26.1 Find an approximation for the equilibrium solution for $\alpha = 1$.

26.2 This is a nonlinear problem. To linearize[a] it around an equilibrium solution, find the Jacobian (or total derivative) J.

26.3 Use `circadian.ipynb` and eigenvalues to check the stability of the equilibrium points for different values of $\alpha \in [0, 100]$.

[a] If you are not comfortable with linearization of a system of ODEs, check exercise 61 on `https://raw.githubusercontent.com/siefkenj/IBLODEs/main/dist/odes.pdf`.

26.1 We get: $x_1 = x_2 = x_3 = x_4 = x_5$ and

$$x_5 = e^{-\alpha x_5}$$

We have to approximate the solutions to this equation, e.g. using Newton's method.

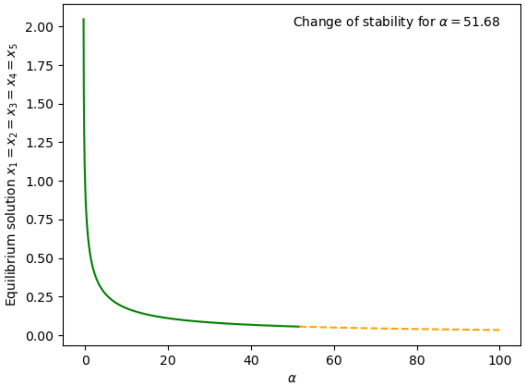
26.2 The Jacobian is:

$$J = \begin{bmatrix} -1 & 0 & 0 & 0 & -\alpha e^{-\alpha x_5} \\ 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix}$$

26.3 The solutions are in `circadian5-sol.ipynb`.

Basically we need to find the (5) eigenvalues for each value of $\alpha \in [0, 100]$ and check when:

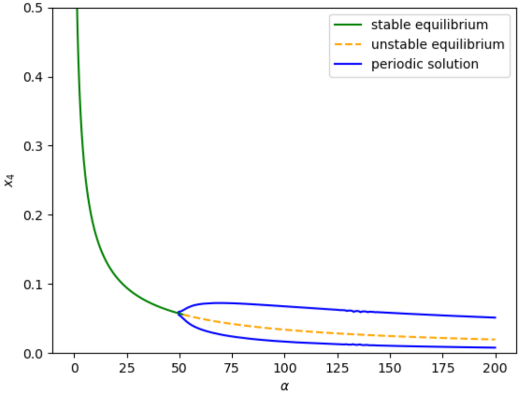
- All negative $\Rightarrow$ stable equilibrium
- One positive $\Rightarrow$ unstable equilibrium



Observe that in this case, we are actually looking for the **unstable** regime, when periodic solutions appear – check the last part of the `Jupyter Notebook`. These only appear when the feedback in strong enough.

Exercise 27

From the previous question, we obtained equilibrium solutions that changed from stable to unstable as we changed the parameter α – see the graph below.



This is called a **bifurcation**.

Another type of bifurcation involves the creation of disappearance of equilibria as a parameter changes.

There are several typical types of bifurcations.

Bifurcations.

A (local) **bifurcation** occurs when a parameter change causes the stability of an equilibrium to change.
We will study four typical types of bifurcations.

1. **Saddle-node bifurcation.** Two equilibria collide and annihilate each other.
2. **Transcritical bifurcation.** An equilibrium exists for all values of a parameter and is never destroyed. However, the equilibrium interchanges its stability with another equilibrium as the parameter changes.
3. **Pitchfork bifurcation.** One equilibrium transitions to three equilibria as a parameter changes.
4. **Hopf bifurcation.** A periodic orbit appears (or disappears) through a change in the stability of an equilibrium point – this means that we transition from purely imaginary to complex eigenvalues.

Decide on the type of bifurcation for each ODE.

27.1 The ODE from Exercise 25.

27.2 The system of ODEs from Exercise 26.

27.3 The ODE $\frac{dx}{dt} = rx - x^2$.

27.4 The ODE $\frac{dx}{dt} = r + x^2$.

27.5 The ODE $\frac{dx}{dt} = rx - x^3$.

27.6 The following system of ODEs as μ changes:

$$\begin{cases} \frac{dx}{dt} &= \mu x - \omega y \\ \frac{dy}{dt} &= \omega x + \mu y \end{cases}$$

27.7 The Lotka-Volterra model for $0 < a < 1$:

$$\begin{cases} \frac{dx}{dt} &= axy - x - 2 + \frac{1}{a} \\ \frac{dy}{dt} &= y - \frac{1}{2}xy - 2 + \frac{1}{a} \end{cases}$$

Exercise 27

27.1 Change of stability bifurcation

27.2 Hopf bifurcation

27.3 Transcritical bifurcation: $x(r-x)=0$ so $x=0$ and $x=r$ are equilibria and they swap stability at $r=0$.

27.4 Saddle-node bifurcation: equilibria only exist for $r<0$, one stable and one unstable.

27.5 Pitchfork bifurcation: $x(r-x^2)=0$ implies

- $r\leq 0$: equilibria at $x=0$
- $r>0$: equilibria at $x=0$ and $x=\pm\sqrt{r}$

See `https://www.desmos.com/calculator/uksexnwbdz` about pitchfork perturbation.

27.6 Hopf bifurcation: Equilibrium at $(0,0)$ and with eigenvalues $\mu\pm\omega i$, so

- $\mu<0$: stable spiral
- $\mu=0$: stable centre (periodic orbit)
- $\mu>0$: unstable spiral

27.7 Equilibrium at $(\frac{1}{a},2)$ and

- $a<1-\frac{\sqrt{3}}{2}\approx 0.134$: two negative eigenvalues (stable)
- $1-\frac{\sqrt{3}}{2}<a<\frac{1}{2}$: stable spiral
- $a=\frac{1}{2}$: stable centre (periodic orbit)
- $a>\frac{1}{2}$: unstable spiral

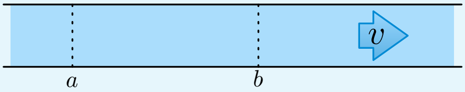
Change in qualitative behaviour at $a=1-\frac{\sqrt{3}}{2}$ and Hopf at $a=\frac{1}{2}$.

Calculations at `bifurcation-LotkaVolterra.i`

Visualize also here `https://www.desmos.com/calculator/aydzcpccy4`

Transport Equation.

Consider a river with the water moving at speed v.



We want to model $w(x,t)$, the density of pollutant in the river at the point x and time t.

28.1 How much pollutant is there in $[a,b]$?

28.2 How does pollutant change in $[a,b]$?

28.3 Find a "conservation of pollutant" equation.

28.4 Simplify the equation to obtain a PDE for $w(x,t)$.

Hint: Recall the FTC: $f(b)-f(a)=\int_a^b f'(x)dx$.

28.1 $T(t) = R\int_a^b w(x,t)\,dx$, where R is the width of the river.

28.2 Pollutant goes in through the left and out through the right, so the change in the amount of pollutant is

$$w(a,t)vR - w(b,t)vR = vR(w(a,t) - w(b,t)).$$

28.3 Because pollutant is neither created or destroyed, we know that:

$$T'(t) = vR(w(a,t) - w(b,t))$$

28.4
$$R\int_a^b \frac{\partial w}{\partial t}(x,t)\,dx = vR(w(a,t) - w(b,t))$$
$$\int_a^b \frac{\partial w}{\partial t}(x,t)\,dx = v(w(a,t) - w(b,t))$$
$$\int_a^b \frac{\partial w}{\partial t}(x,t)\,dx = -v\int_a^b \frac{\partial w}{\partial x}(x,t)\,dx$$

$$\int_a^b \frac{\partial w}{\partial t} + v\frac{\partial w}{\partial x}\,dx = 0$$

Because a, b are arbitrary points in the river, we can conclude that

$$\frac{\partial w}{\partial t} + v\frac{\partial w}{\partial x} = 0 \qquad \text{or} \qquad w_t + v\cdot w_x = 0$$

Here is a Jupyter notebook with the Lax-Friedrichs Method approximation for the transport equation:

- `transport_LaxFriedrichs.ipynb`

Method of Characteristics.

This is a method to solve a specific type of Partial Differential Equations:

$$u_t(x,t) + f(x,t) \cdot u_x(x,t) = g(x,t).$$

The idea is to interpret the left-hand side as a total derivative with respect to t:

$$\frac{du}{dt}(x(t),t) = u_t(x,t) + f(x,t)u_x(x,t),$$

which implies that

$$\begin{cases} \frac{dx}{dt} = f(x,t) & \text{(moving observer)} \\ \frac{d}{dt}[u(x(t),t)] = g(x(t),t) & \text{(solution for the observer)} \end{cases}$$

The moving observers $x(t)$ are called the *characteristics*. This method allows us to "transform" a PDE into two ODEs.

Video: `https://youtu.be/tNP286WZw3o`

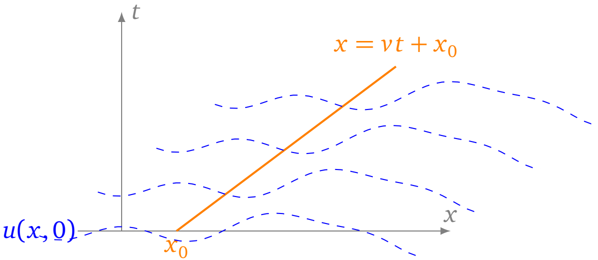
29.1 Find the solution of the transport equation from Exercise 28 using the Method of Characteristics with the initial condition $u(x,0) = p(x)$.

29.2 Find the solution of the same problem with an accelerating river: $v = 3t^2$.

29.3 Find the solution for $w_t + 5w_x = 2$.

29.4 Find the solution for $w_t + 3t^2 w_x = 2$.

29.5 Find the solution for $w_t + 3t^2 w_x = -x$.



29.1 We need to solve

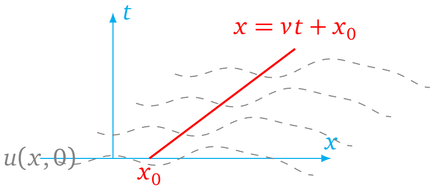
- $\frac{dx}{dt} = v$ $\rightarrow$ an observer moving along the river at the same speed as the river
- $\frac{du}{dt} = 0$ $\rightarrow$ for such an observer looking at the river, the pollutant density doesn't change

$$\begin{cases} x = vt + x_0 \\ u(x(t), t) = C \end{cases}$$

This means that when $t = 0$, we get $u(x_0, 0) = C = p(x_0)$, and $x_0 = x - vt$, so

$$u(x,t) = C = p(x_0) = p(x - vt).$$

The idea in a graph:



Here we can run an approximation for a specific $u(x, 0)$ and $v = -1.2$:

- `transport_LaxFriedrichs.ipynb`

29.2 The idea is similar but we get

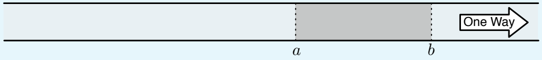
$$\begin{cases} x = t^3 + x_0 \\ u(x(t), t) = C \end{cases}$$

This means that when $t = 0$, we get $u(x_0, 0) = C = p(x_0)$, and $x_0 = x - t^3$, so

$$u(x,t) = C = p(x_0) = p(x - t^3).$$

Exercise 30

Traffic Flow.



We want to model how traffic flows on a one way road.

Let

- $\rho(x,t) =$ density of cars (number of cars per km)
- $\phi(x,t) =$ number of cars passing the point x per hour

And assume:

(C) Cars are conserved (they are not destroyed nor created on the road)

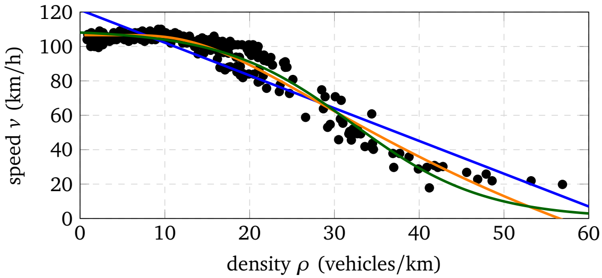
30.1 What is the total number of cars in the section of the road $x \in [a,b]$ at time t?

30.2 How does the total number of cars change in $[a,b]$?

30.3 Obtain an equation relating $\rho(x,t)$ and $\phi(x,t)$. The equation should not include a or b.

We need to model how fast cars move on the road: $\phi(x,t)$.

Below we graphed measurements for density and speed at the highway 401[a] together with three different models to fit the data.



It shows three models:

- Greenshields model (linear fit of the data): $v(\rho) = v_{\max}\left(1 - \dfrac{\rho}{\rho_{\max}}\right)$
- Newel model
- Logistic model: $v(\rho) = v_{\max}/\left(1 + e^{-k(\rho-\rho_0)}\right)$

30.4 Using the Greenshields model, find an expression for $\phi(x,t)$.

30.5 Obtain a PDE for $\rho(x,t)$.

[a] Data from the paper "Calibrating Steady-State Traffic Stream and Car-Following Models Using Loop Detector Data" by H Rakha and M Arafeh

30.1 $C(t) = \int_a^b \rho(x,t)\,dx$

30.2 $C'(t) = \phi(a,t) - \phi(b,t)$

30.3 From the previous two parts, we get

$$\int_a^b \rho_t\,dx = \phi(a,t) - \phi(b,t)$$

By using the FTC, we have $\phi(b,t) - \phi(a,t) = \int_a^b \phi_x(x,t)\,dx$, so we conclude that

$$\int_a^b \rho_t + \phi_x(x,t)\,dx = 0$$

Because this integral must be true for any values of a, b, we conclude that the integrand must be zero:

$$\rho_t + \phi_x(x,t) = 0$$

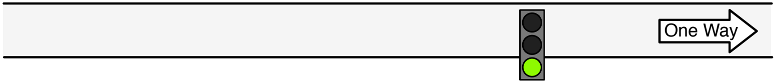
30.4 $\phi(x,t) = \rho v(\rho)$

30.5 We can expand the equation we found before:

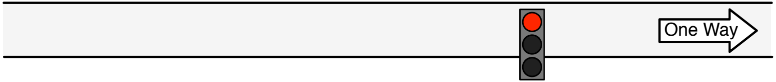
$$\begin{aligned} \rho_t + \phi_x(x,t) &= 0 \\ \rho_t + \rho_x v_{\max}\left(1 - \frac{\rho}{\rho_{\max}}\right) - \rho v_{\max}\frac{\rho_x}{\rho_{\max}} &= 0 \\ \rho_t + \rho_x v_{\max}\left(1 - \frac{2\rho}{\rho_{\max}}\right) &= 0 \end{aligned}$$

Exercise 31

Let us study two interesting cases.



31.1 What is the initial car density $\rho(x,0) = \rho_0(x)$ on a one way road with a traffic light that just turned from **red** to **green**?



31.2 To model a light turning from **green** to **red**, we need to be more creative. What is an initial car density $\rho(x,0) = \rho_0(x)$ that will guarantee incoming cars have to stop at the red light?

Traffic flow scenario.
We want to solve the following traffic flow problem:

$$\rho_t + v_{\max}\left(1 - \frac{2\rho}{\rho_{\max}}\right)\rho_x = 0 \qquad \text{(Traffic flow model)}$$

$$r(x,0) = f(x) = \begin{cases} \rho_{\min} & \text{for } x < 0 \\ \rho_{\max} & \text{for } x > 0 \end{cases}$$

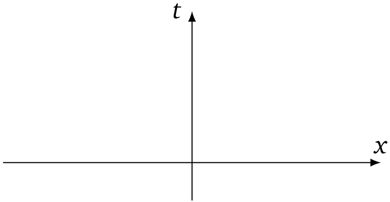
Consider the following parameters:

- $v_{\max} = 60$
- $\rho_{\max} = 120$
- $\rho_{\min} = 20$

32.1 What are the moving observers (characteristics) $x(t)$ for this problem?

32.2 What is the density $\rho(x,t)$?

32.3 Sketch the characteristics and mark the values of ρ on the same graph below.



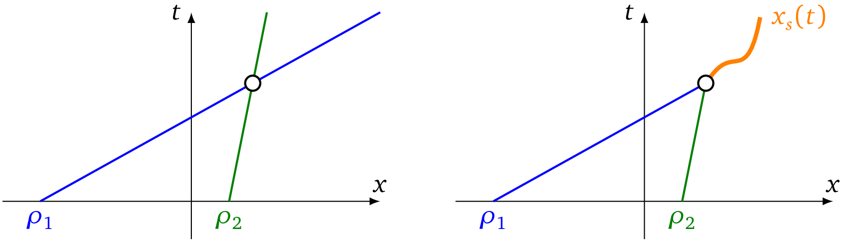
32.4 What is $\rho(0,2)$?

Exercise 33

When characteristics intersect, this means that the solution **cannot be continuous**.

So we need to find a **discontinuous** solution.

- Assume that the discontinuity forms a curve $x_s(t)$.



33.1 What should the discontinuity $x_s(t)$ be?

Exercise 34

We need to step back for a moment and review some Calculus.

Consider a function

$$F(x) = \int_0^z g(x)\,dx$$

and consider a differentiable function $h(t)$.

34.1 What is $F'(z)$?

34.2 What is $F'(g(t))$? What is $[F(h(t))]'$?

34.3 What is $\left[\int_0^{h(t)} g(x)\,dx\right]'$? What is $\left[\int_{h(t)}^{1} g(x)\,dx\right]'$?

We need to go back to the derivation of the traffic flow model.

We had the following:

$$\frac{d}{dt}\left[\int_a^b \rho(x,t)\,dx\right] = \phi(a,t) - \phi(b,t)$$

We then took the derivative inside the integral, because we assumed that the density ρ was differentiable (thus continuous). Now we know it is not, so we must break up the interval of integration into "chunks" where ρ is continuous.

We now assume that $\rho(x,t)$ is discontinuous across $x = x_s(t)$.

35.1 Expand the left-hand side of the equation into integrals with continuous integrands.

35.2 We know that $\phi = \phi(\rho)$. Take the limits

$$a \to (x_s(t))^- \quad \text{and} \quad b \to (x_s(t))^+$$

and obtain an ODE for $x_s(t)$.

This ODE is called the **Rankine-Hugoniot shockwave condition**.

35.1 We have

$$\frac{d}{dt}\left[\int_a^b \rho(x,t)\,dx\right] = \frac{d}{dt}\left[\int_a^{x_s(t)} \rho(x,t)\,dx + \int_{x_s(t)}^b \rho(x,t)\,dx\right]$$

Using the previous exercise, we get

$$\left[\rho(x_s^-(t),t) - \rho(x_s^-(t),t)\right] x_s'(t) + \int_a^b \rho_t(x,t)\,dx$$

35.2 Let us define the following

- $\rho^-(t) = \lim_{x\to x_s^-(t)} \rho(x,t)$
- $\rho^+(t) = \lim_{x\to x_s^+(t)} \rho(x,t)$

So when we take the limits, we obtain

$$(\rho^-(t) - \rho^+(t))x_s'(t) = \phi(\rho^-(t)) - \phi(\rho^+(t))$$

We get

$$x_s'(t) = \frac{\phi(\rho^-) - \phi(\rho^+)}{\rho^- - \rho^+}$$

This condition is called the **Rankine-Hugoniot shockwave condition**.

Exercise 36

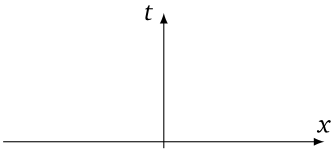
36.1 Use the Rankine-Hugoniot shockwave condition to find the full solution of Exercise 32.

36.2 Compare the solution to the numerical solution using the Lax-Friedrichs method in `traffic_flow_LaxFriedrichs.ipynb`.

Note that to use this method, we wrote the PDE as $\rho_t + (\phi(\rho))_x = 0$ with $\phi(\rho) = v_{\max}\left(1 - \frac{\rho}{\rho_{\max}}\right)\rho$.

Note also that the method is very sensitive to the choice of Δx and Δt: it only works when $\frac{\Delta t}{\Delta x}$ is small enough.

36.3 Trace the paths of the cars starting at $x_0 = -10, -5, 0, 2.5$.



36.4 What happens when cars slow down gradually? Find the solution for the initial condition

$$r(x,0) = f(x) = \begin{cases} \rho_{\min} & \text{for } x < -1 \\ \rho_{\max} + (\rho_{\min} - \rho_{\max})x & \text{for } -1 \le x \le 0 \\ \rho_{\max} & \text{for } x > 0 \end{cases}$$

36.5 How would the model change if there is an on-ramp at $x = 0$?

36.1 The Rankine-Hugoniot condition yields:

$$x_s'(t) = \frac{\phi(20) - \phi(120)}{20 - 120} = -\frac{1000 - 0}{100} = -10$$

where we recall that $\phi(\rho) = v_{\max}\left(1 - \frac{\rho}{\rho_{\max}}\right)\rho = 60\left(1 - \frac{\rho}{120}\right)\rho.$

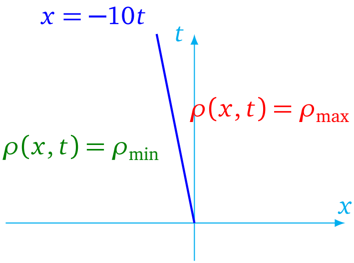
We also know that the discontinuity starts at the point $(x,t) = (0,0)$, so

$$x_s(t) = -10t$$

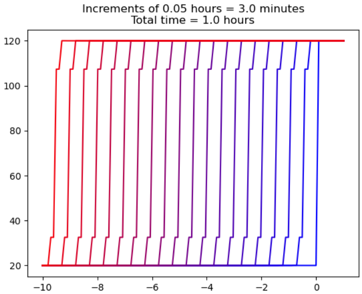
This means that the solution is

$$\rho(x,t) = \begin{cases} \rho_{\min} & \text{for } x < x_s(t) \\ \rho_{\max} & \text{for } x > x_s(t) \end{cases}$$

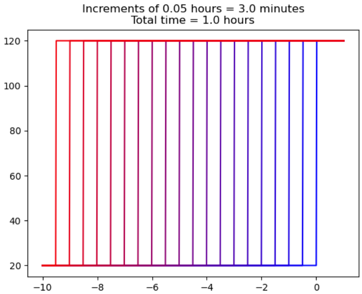
In practice this means that cars are accumulating behind the traffic sign ($\rho_{\max}$) means cars are stopped. The cars are accumulating at the speed of 10 km/h.



36.2 When we run the numerical solution (click here to see an animation), we get the following:



larger Δx and $\frac{\Delta t}{\Delta x} = 0.02$



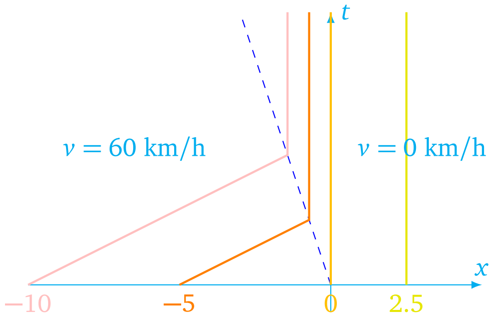
smaller Δx and $\frac{\Delta t}{\Delta x} = 0.1$

If the resolution in x is not good enough, then we see some artifacts from the numerical approximation.

We can estimate the speed of the shockwave: it takes 4 time-steps to get to from $x = 0$ to $x = -2$:

- Speed of the shockwave $= \frac{2}{4 \cdot (0.05)} = 10$ km/h.

36.3



Here is an animation of the solution: traffic_flow-animation.mp4

36.4 In this case, the shockwave doesn't start at $t = 0$, but when the characteristics meet for the first time.

Probability Models

Introduction to continuous probability.

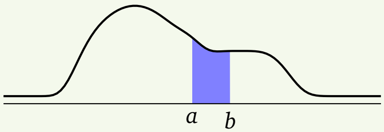
- Continuous random variables take values according to a **probability density function (pdf)**
- We evaluate the probability that a random variable has a value in an interval by integrating the pdf over that interval:

$$\Pr(a \le T \le b) = \int_a^b f_T(t)\,dt,$$

where $f_T(t)$ is the pdf.

- Since T must have a value, we have

$$\int_{-\infty}^{\infty} f_T(t)\,dt = 1.$$



- It is convenient to define the **cumulative distribution function (cdf)** $F_T(t)$ and the complementary cumulative distribution function (ccdf) $\tilde{F}_T(t)$ as

$$F_T(t) = \Pr(T \le t) = \int_{-\infty}^{t} f_T(\tau)\,d\tau$$
$$\tilde{F}_T(t) = \Pr(T \ge t) = \int_{t}^{+\infty} f_T(\tau)\,d\tau$$

- These functions are related by $F_T(t) = 1 - \tilde{F}_T(t)$.
- By the FTC, we also have $\frac{dF_T}{dt}(t) = f_T(t)$.
- The **mean** of the random variable T is

$$\mu_T = \int_{-\infty}^{\infty} t \cdot f_T(t)\,dt.$$

Conditional Probability. $\Pr(A|B) =$ Probability that A will happen given that B has happened $= \dfrac{\Pr(A \text{ and } B)}{\Pr(B)}$.

Hazard function.

We also define the *hazard function* $h_T(t)$, which is the probability density function of an event that has not happened yet will occur imminently.
We then have

$$\begin{aligned}
h_T(t) &= \lim_{\Delta t\to 0^+} \frac{\Pr(t \le T \le t+\Delta t | T \ge t)}{\Delta t}\\
&= \lim_{\Delta t\to 0^+} \frac{\Pr(t \le T \le t+\Delta t \text{ and } T \ge t)}{\Pr(T\ge t)\Delta t}\\
&= \lim_{\Delta t\to 0^+} \frac{\Pr(T \le t+\Delta t) - \Pr(T \le t)}{\Pr(T\ge t)\Delta t}\\
&= \underbrace{\lim_{\Delta t\to 0^+} \frac{F_T(t+\Delta t) - F_T(t)}{\Delta t}}_{=\frac{dF_T}{dt}} \cdot \frac{1}{\tilde{F}(t)}\\
&= \frac{f_T(t)}{\tilde{F}(t)}
\end{aligned}$$

38.1 Show that $\frac{d\tilde{F}_T}{dt} = -f_T$.

38.2 Find an ODE that relates $\tilde{F}_T$ and $h(t)$ and use it to show that

$$\tilde{F}_T(t) = e^{-\int_{-\infty}^{t} h_T(\tau)\, d\tau}$$

38.3 Show that the mean satisfies

$$\mu_T = \int_{-\infty}^{\infty} \tilde{F}_T(t)\, dt.$$

Exponential Random Variable.
The exponential random variable $T \sim \mathrm{Exp}(\gamma)$ is defined by the pdf:

$$f_T(t) = \begin{cases} \gamma e^{-\gamma t} & \text{if } t \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Let us study the **exponential random variable**:

$$T \sim \mathrm{Exp}(\gamma).$$

39.1 Find an expression for its cdf $F_T(t)$.

39.2 Find an expression for its ccdf $\tilde{F}_T(t)$.

39.3 Find an expression for its hazard function $h_T(t)$.

39.4 What is its mean μ_T?

The hazard function is a constant, so knowing that the random variable is yet to occur, doesn't us give any information.

Memoryless Random Variable.
A random variable is said to be **memoryless** if

$$\Pr(T > t + s | T > t) = \Pr(T > s).$$

39.5 Show that the Exponential is memoryless.

Exercise 39

39.1 $F_T(t) = \int_{-\infty}^{t} f_T(\tau)\, d\tau = \int_0^t \gamma e^{-\gamma\tau}\, d\tau = 1 - e^{-\gamma t}$

39.2 $\tilde{F}_T(t) = 1 - F_T(t) = e^{-\gamma t}$.

39.3 $h_T(t) = \dfrac{f_T(t)}{\tilde{F}_T(t)} = \gamma$.

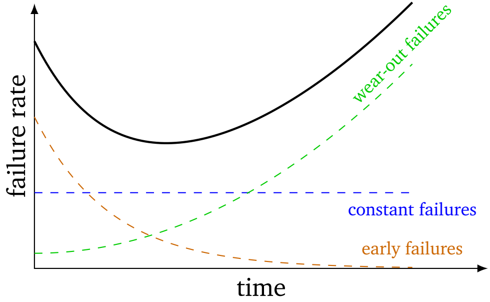
39.4 $\mu_T = \int_0^{\infty} \tau f_T(\tau)\, d\tau = \int_0^{\infty} \tilde{F}_T(\tau)\, d\tau = -\dfrac{1}{\gamma} e^{-\gamma\tau}\big|_0^{\infty} = \dfrac{1}{\gamma}$

39.5 $\Pr(T > t + s | T > t) = \dfrac{\Pr(T > t + s)}{\Pr(T > t)} = \dfrac{e^{-\gamma(t+s)}}{e^{-\gamma t}} = e^{-\gamma s} = \Pr(T > s)$.

Exercise 40

Backblaze is a cloud storage company that uses thousands of hard drives.[a]

The failure rate for a hard drive, also known as hazard rate, is composed of three terms:



[a]They publish data on their hard drives `https://www.backblaze.com/b2/hard-drive-test-data.html`.

When we add the three terms, we get a *bathtub* curve.

Should Backblaze always use the hard drives with the longest mean time between failures (MTBF)?

40.1 Suppose that there are N hard drive types, each with a known fixed cost c_n, a known profit p_n per unit time that hard drive is operational, and a random lifetime T_n with a known distribution. Backblaze has limited storage, so they can only have M hard drives. If they use x_n drives of type n, what is their profit function?

40.2 What is their expected profit?

40.3 If they choose to maximize the expected profit, what is the solution?

This means that they will *put all their eggs in one basket*!

This is a risky strategy. Especially if hard drives with long-lives have a large variance in those lives.

Backblaze Q2 2024 Annualized Failure Rates

Drive models as of 6/30/2024 with > 100 drives and > 10,000 drive days for Q2 2024

MFR	Model	Size (TB)	Drive Count	Avg. Age (Months)	Drive Days	Drive Failures	Q2 2024 AFR
WDC	WUH722222ALE6L4	22	13,140	3.6	878,867	33	1.37%
Seagate	ST16000NM002J	16	462	19.7	41,202	-	0.00%
WDC	WUH721816ALE6L4	16	26,467	15.3	2,407,414	26	0.39%
Seagate	ST16000NM001G	16	33,411	22.8	2,995,255	68	0.83%
Toshiba	MG08ACA16TA	16	38,803	13.7	3,517,575	114	1.18%
Toshiba	MG08ACA16TE	16	5,936	32.4	540,962	24	1.62%
Toshiba	MG08ACA16TEY	16	5,202	30.8	464,010	22	1.73%
WDC	WUH721816ALE6L0	16	3,028	30.1	270,162	14	1.89%
Seagate	ST14000NM000J	14	139	6.3	11,635	-	0.00%
Toshiba	MG07ACA14TEY	14	710	32.1	63,896	1	0.57%
WDC	WUH721414ALE6L4	14	8,486	42.2	772,171	20	0.95%
Toshiba	MG07ACA14TA	14	37,891	43.7	3,451,376	105	1.11%
Seagate	ST14000NM001G	14	10,670	40.3	972,122	45	1.69%
Seagate	ST14000NM0138	14	1,363	42.9	124,708	16	4.68%
Seagate	ST12000NM001G	12	13,183	41.0	1,197,765	39	1.19%
HGST	HUH721212ALE600	12	2,580	56.2	232,648	11	1.73%
Seagate	ST12000NM0008	12	19,300	50.4	1,758,710	122	2.53%
HGST	HUH721212ALE604	12	13,115	38.8	1,195,749	94	2.87%
Seagate	ST12000NM000J	12	473	5.2	35,204	4	4.15%
HGST	HUH721212ALN604	12	10,496	60.5	957,573	188	7.17%
Seagate	ST12000NM0007	12	1,096	55.3	101,292	33	11.89%
Seagate	ST10000NM0086	10	1,073	78.4	98,280	12	4.46%
HGST	HUH728080ALE600	8	1,097	74.0	100,003	3	1.09%
Seagate	ST8000NM000A	8	247	17.1	21,380	1	1.71%
Seagate	ST8000DM002	8	9,173	92.6	836,716	41	1.79%
Seagate	ST8000NM0055	8	13,726	80.9	1,255,225	95	2.76%
HGST	HMS5C4040BLE640	4	9,776	92.2	898,707	7	0.28%
HGST	HMS5C4040ALE640	4	767	90.4	83,549	1	0.44%
Seagate	ST4000DM000	4	2,576	99.5	292,534	58	7.24%
		Totals	**284,386**		**25,576,690**	**1,197**	**1.71%**

Backblaze

To account for this, we can add a constraint that the variance of the profit σ_P^2 not be too large. This will result in a more diverse selection of hard drives. Unfortunately, for most hard drive lifetime distributions, this constraint is non-linear in the decision variables.

Exercise 41

Poisson Process.

The most famous stochastic process is the **Poisson Process** to count events.

It assumes that:

- The events are randomly distributed
- It is memoryless: all t's are equal. The probability of an event happening doesn't depend on t
- Events cannot un-occur, so counting only increases.

Let $N(t)$ be the (random) number of events that have occurred up to time t. Because events are memoryless, this means that the random times between events T are **exponentially distributed** with pdf

$$f_T(t) = \gamma e^{-\gamma t}.$$

Has we have seen in exercise 39:

- The cdf is $F_T(t) = 1 - e^{-\gamma t}$ and the ccdf is $\tilde{F}_T(t) = e^{-\gamma t}$
- The hazard function is $h_T(t) = \gamma$ a constant.

Start with $N(0) = 0$ and define $p_n(t) = \Pr(N(t) = n)$, the probability that n events have occurred up to time t.

41.1 What is $p_n(0)$?

41.2 Take $n = 0$. Explain in words why $p_0(t) = \tilde{F}_T(t)$.

41.3 A different approach:

$$\underbrace{\frac{dp_0}{dt}(t)}_{\substack{\text{change in prob} \\ \text{that } N(t)=0}} = + \underbrace{0}_{\text{can never increase}} - \underbrace{p_0(t)h_T(t)}_{\substack{\text{decreases if } N \text{ was 0} \\ \text{and an event occurs}}}$$

Solve this differential equation to find $p_0(t)$.

41.4 For $n > 0$, we can find an ODE in a similar way. The change in probability that $N(t) = n$:

(a) decreases if ...

(b) increases if ...

41.5 Obtain an ODE and show that $p_n(t) = \frac{(\gamma t)^n}{n!} e^{-\gamma t}$ is a solution.

41.6 Show that it is normalized, i.e. $\sum_{n=0}^{\infty} p_n(t) = 1$.

41.7 Calculate its mean $\sum_{n=0}^{\infty} n p_n(t)$.

41.1 $p_n(0) = \delta_{n,0}$

41.2 It is the probability that no events have occurred until time t, which is the complement of the probability of an event happening between time 0 and t, so $p_0(t) = 1 - F_T(t) = \tilde{F}_T(t) = e^{-\gamma t}$.

41.3 The ODE is $\frac{dp_0}{dt} = -\gamma p_0$, so the solution is $p_0(t) = e^{-\gamma t}$.

41.4 (a) increases if N was $n-1$ and an event occurs imminently

(b) decreases if N was n and an event occurs imminently

41.5

$$\underbrace{\frac{dp_n}{dt}(t)}_{\substack{\text{change in prob} \\ \text{that } N(t)=n}} = + \underbrace{p_{n-1}(t)h_T(t)}_{\substack{\text{increases if } N \text{ was } n-1 \\ \text{and an event occurs}}} - \underbrace{p_n(t)h_T(t)}_{\substack{\text{decreases if } N \text{ was } n \\ \text{and an event occurs}}}$$

This gives

$$\frac{dp_n}{dt}(t) = \gamma p_{n-1}(t) - \gamma p_n(t)$$

41.6 The series

$$\sum_n \frac{(\gamma t)^n}{n!}$$

is the Taylor series for $e^{\gamma t}$.

41.7 Consider $g_n(t) = \frac{(\gamma t)^n}{n!}$. Then

$$t g_n'(t) = n g_n(t)$$

So

$$\sum_n n \frac{(\gamma t)^n}{n!} = t \left(\sum_n \frac{(\gamma t)^n}{n!} \right)' = t \left(e^{\gamma t} \right)' = t \gamma e^{\gamma t}$$

Thus the mean is $E(N(t)) = \gamma t$, which is typically denoted by λ.

Poisson Process.

From the exercise 41, we usually denote $\lambda = \gamma t$ and so we have:

$$\Pr(N = k) = \frac{\lambda^k}{k!} e^{-\lambda},$$

Its mean is λ and so is the variance.

You can watch these two videos to get a better idea of what is $p_n(t)$: `https://youtu.be/3z-M6sbGIZ0` and `https://youtu.be/Jkr4FSrNEVY`

Exercise 42

Aquarium Problem[a]**.**
A pet store sells large aquariums. They sell approximately **one aquarium per week** (call this the parameter λ).
Not wanting to maintain too much stock (the aquariums are large and fragile), the store orders 3 aquariums at the end of the week if they are completely out of stock.
How missed sales result from this policy?

[a] based on a problem from Meerschaert's 'Mathematical Modeling'.

Let

- X_n = number of aquariums in stock at the beginning of week n
- D_n = number of aquariums demanded in week n

Any available aquariums that are demanded are purchased and assume that when the store orders aquariums, they arrive right away.

42.1 Why is X_n random?

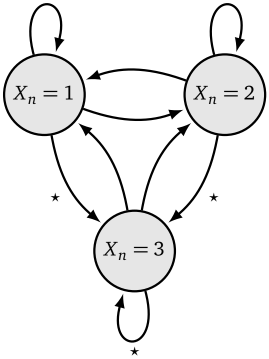
42.2 Assume that D_n is Poisson distributed. Why is this reasonable?

42.3 Then what is $\Pr(D_n = k)$?

42.4 There are only 3 states of D_n: 1, 2, 3.

The following diagram shows the possible changes in stock from one week to the next.

Complete the following diagram with the probability of transitioning from one state to another for all the arrows except the ones marked with $\star$.



42.5 For the remaining arrows, marked with $\star$, observe that if there is one aquarium in inventory and 3 clients come, then the store will only be able to sell one of them and order new aquariums, so at the beginning of next week there will be 3 aquariums. Because of this, complete the remaining $\star$ arrows by using the complementary probability.

42.6 Create a transition matrix $\mathcal{P}$ with

- P_{ij} = probability of transitioning from state i to j

42.7 Let

$$\vec{\pi}_n = \begin{bmatrix} \Pr(X_n = 1) \\ \Pr(X_n = 2) \\ \Pr(X_n = 3) \end{bmatrix}$$

Find a relation between $\vec{\pi}_{n+1}$ and $\vec{\pi}_n$.

42.8 Use `aquarium.ipynb` to approximate the solution for the long term states. Interpret the results.

42.9 The original question was about how much business is the store missing. Write the expected percentage of weeks when the store lost at least one sale as a probability.

42.10 Calculate W:

$$W = \lim_{n\to\infty} \Pr(D_n > X_n)$$
$$= \lim_{n\to\infty} \sum_{i=1}^{3} \Pr(D_n > X_n | X_n = i) \Pr(X_n = i),$$

using the limiting vector $\vec{\pi}$ found. Interpret the result.

42.11 We can calculate the average lost sales too by considering:

$$L = \lim_{n\to\infty} \sum_{i=1}^{3} (D_n - i) \Pr(D_n > X_n | X_n = i) \Pr(X_n = i).$$

Calculate L, the expected number of lost aquarium sales. Interpret the results.

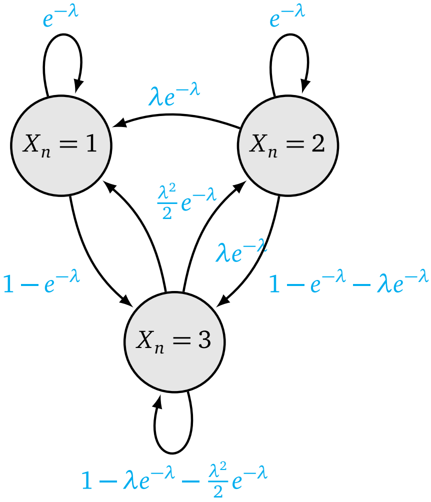
42.12 What is $S(L^\star, \lambda)$?

42.1 Because D_n is random.

42.2 Because the arrival of customers can happen at any time, they are independent of each other, and they don't depend on t (in reality they do), but because we are measuring in weeks, it's ok.

42.3 $\Pr(D_n = k) = \frac{\lambda^k}{k!} e^{-\lambda}$.

42.4



42.6

$$\mathcal{P} = \begin{bmatrix} e^{-\lambda} & 0 & 1-e^{-\lambda} \\ \lambda e^{-\lambda} & e^{-\lambda} & 1-e^{-\lambda}-\lambda e^{-\lambda} \\ \frac{\lambda^2}{2}e^{-\lambda} & \lambda e^{-\lambda} & 1-\lambda e^{-\lambda}-\frac{\lambda^2}{2}e^{-\lambda} \end{bmatrix}$$

42.7 We have $\vec{\pi}_{n+1} = \vec{\pi}_n \mathcal{P} \quad \Leftrightarrow \quad \vec{\pi}_{n+1} = \mathcal{P}^T \vec{\pi}_n$.

42.8 The long term result is: $\vec{\pi} = \begin{bmatrix} 0.28471134 \\ 0.26313999 \\ 0.45214867 \end{bmatrix}$.

This means that with $\lambda = 1$, that is with 1 expected customer per week, the store should expect to have :

- 1 aquarium in inventory 28.5% of the weeks
- 2 aquariums in inventory 26.3% of the weeks
- 3 aquariums in inventory 45.2% of the weeks

This means that about half of the time, they have a full inventory. Perhaps this is due to the fact that they ran out of aquariums and just ordered more, so they might be losing a lot of sales when they run out of stock.

42.9 It is the probability that the demand is higher than the inventory: $\Pr(D_n > X_n)$.

42.10 W is the long term expected percentage of weeks when sales are lost:

$$\begin{aligned} W &= \lim_{n\to\infty} \sum_{i=1}^{3} \Pr(D_n > X_n | X_n = i) \Pr(X_n = i) \\ &= \lim_{n\to\infty} \sum_{i=1}^{3} \Pr(X_n = i) \sum_{j=i+1}^{\infty} \Pr(D_n = j) \\ &= \sum_{i=1}^{3} \pi_i \sum_{j=i+1}^{\infty} \frac{\lambda^j}{j!} e^{-\lambda} \\ &= e^{-\lambda} \left[\pi_1 \sum_{j=2}^{\infty} \frac{\lambda^j}{j!} + \pi_2 \sum_{j=3}^{\infty} \frac{\lambda^j}{j!} + \pi_3 \sum_{j=4}^{\infty} \frac{\lambda^j}{j!} \right] \end{aligned}$$

The series in j is the tail of the Taylor series for the exponential e^{λ}, so we can write it as

$$\begin{aligned} W =& e^{-\lambda} \Big[\pi_1 (e^{\lambda} - 1 - \lambda) + \pi_2 (e^{\lambda} - 1 - \lambda - \tfrac{\lambda^2}{2}) \\ &+ \pi_3 (e^{\lambda} - 1 - \lambda - \tfrac{\lambda^2}{2} - \tfrac{\lambda^3}{6}) \Big] \end{aligned}$$

Using `aquarium-sol.ipynb`, we get $W = 0.105$, which means that the store loses sales about 11% of the weeks.

42.11 Similarly, we can calculate $L = 0.14256$, so the store should expect to lose an average of 0.14 sales per week. Given that the expected number of sales is 1 per week, this is about 14% of their business and it might be worthwhile to re-evaluate their stocking policy.

42.12 To calculate the sensitivity, we need the approximate formula:

$$S(L,\lambda) \approx \frac{L(1+\Delta) - L(1)}{\Delta} \cdot \frac{1}{L(1)} = 1.859$$

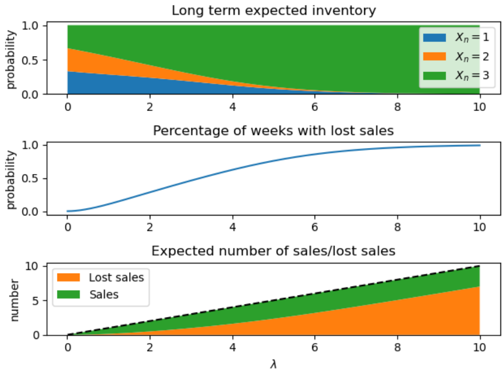
Which is just moderately sensitive to λ.

I also calculated the expected number of sales (should be less than λ since we are losing some sales):

$$\begin{aligned}
E &= \sum_{i=1}^{3}\sum_{j=1}^{i} jP(D_n = j)P(X_n = i) + \sum_{i=1}^{3}\sum_{j=i+1}^{\infty} iP(D_n = j)P(X_n = i) \\
&= \sum_{i=1}^{3}\sum_{j=1}^{i} j\frac{\lambda^j}{j!}e^{-\lambda}\pi_i + \sum_{i=1}^{3}\sum_{j=i+1}^{\infty} i\frac{\lambda^j}{j!}e^{-\lambda}\pi_i \\
&= e^{-\lambda}\left[\sum_{i=1}^{3}\pi_i\lambda\sum_{j=0}^{i-1}\frac{\lambda^j}{j!}\pi_i + \sum_{i=1}^{3} i\pi_i \sum_{j=i+1}^{\infty} i\frac{\lambda^j}{j!}\right] \\
&= 0.857 \quad \text{for } \lambda = 1.
\end{aligned}$$

So the store only makes about 85.7% of possible sales.

I ran the code for values of $\lambda \in (0,10]$ to get the plots:



Observe how the number of sales converges to 3 as $\lambda \to \infty$, which makes sense since that is the maximum inventory. The more clients that come to the store, the more likely to sell all the inventory every week.

Exercise 43

Construct a Markov chain model for the heat equation interpreting it as particles randomly bumping around and averaging their temperature when they meet.

Stochastic Simulation. It is very rare to be able to obtain analytic results for probabilistic models, so we will study simulating them.

Discrete Event Method.
Also known as the Gillespie or Stochastic simulation algorithm.
In this method, pseudo-random numbers determine the times of each event and the simulation progresses by jumping to the time of the first event, updating the state according to the event, sampling new times for the events, and repeating. This produces an exact realization of the stochastic process.
This method will be slow if there are lots of events.

τ-leaping Method.
Similarly to Euler's method for ODEs, we take a fixed time step Δt (usually it is denoted by τ) and decide independently how many of each event occurred. This is an approximation in that as events occur during the finite time step, the rates should have changed, but were not.
This method will be slow if many steps do not result in any events.

Here is some code simulating an Exponential Distribution using each method.

- `numerical-stochastic.ipynb`

Exercise 45

Recall Exercise 42.

45.1 Using the Jupyter Notebook `aquarium-MonteCarlo.ipynb`, simulate the aquarium problem.

45.2 The store is considering an alternative re-stocking policy: when the stock is down to 1 aquarium, they flip a coin and decide whether to re-stock to 3 aquariums or not.

Simulate this new store policy in the same Jupyter Notebook and compare the results.

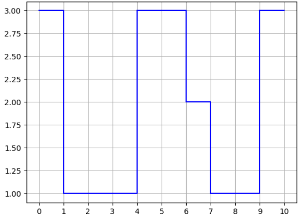
45.3 Run the last part of the Jupyter Notebook to combine the results of these two policies. You need to add titles and legends to the graphs, and labels to the $x-$ and $y-$ axes. Compare the results.

45.4 Create an economic model for the store that includes: profit from each aquarium sale, stocking cost per aquarium in stock, shipping cost when re-ordering new aquariums.

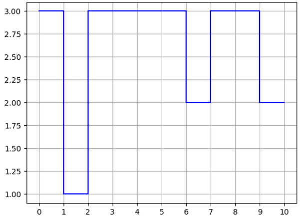
Evaluate which of the previous two policies is better.

Solution python: `aquarium-MonteCarlo-sol.ipynb`

45.1 Plot one simulation for the original re-stocking policy.



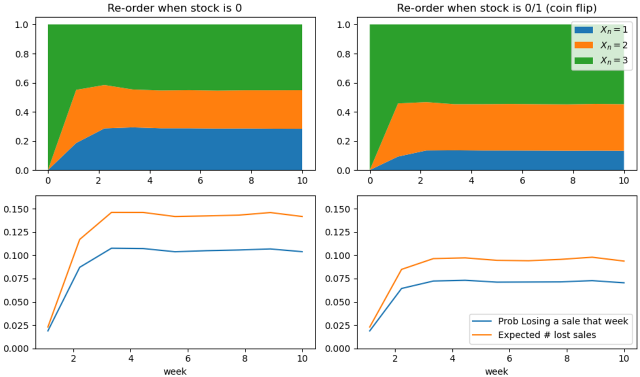
45.2 Plot the same simulation for the modified re-stocking policy.



Observations:

- Week 1: the modified store decided not to order new aquariums, even though they only had 1 in stock
- Week 2: the modified store re-stocked
- Week 3: the original store had 1 aquarium in stock and sold it, then re-stocked
- Week 3: the modified store, had 3 aquariums in stock and sold at least 2, because they also re-stocked
- Week 3: This means that the original store lost 1 or 2 more sales than the modified store.

45.3 From the plots below, we can see that the modified store loses much fewer sales.



Exercise 46

Recall the SIR model:

$$\frac{dS}{dt} = -\beta SI$$
$$\frac{dI}{dt} = \beta SI - \gamma I$$
$$\frac{dR}{dt} = \gamma I$$

We now want to create a stochastic simulation for the same model:

- A *susceptible* individual becomes *infected* with probability βI per unit time;
- An *infected* individual becomes *recovered* with probability γ per unit time.

An *event* happens when one of these two changes of circumstances happens and we assume that the time between events is *memoryless*, hence it should be *exponentially* distributed with the appropriate rate.

Since the number of infected individuals changes with every event, the rate will need to be updated with each event.

46.1 Explain why the *discrete event* method will give a better approximation than the τ-*leaping* method.

46.2 We are expecting lots of events, so we will use the τ-leaping method. we use the binomial distribution (the discrete version of the Poisson distribution) to determine whether an individual became infected or not. Then complete the following for **one time step** of length Δt:

- Pr(new infectives$=k$) $\quad =$
- Pr(new recovereds$=k$) $\quad =$

46.3 Create a simulation for $t \in [0, 10]$ and $\Delta t = 0.1$ with:

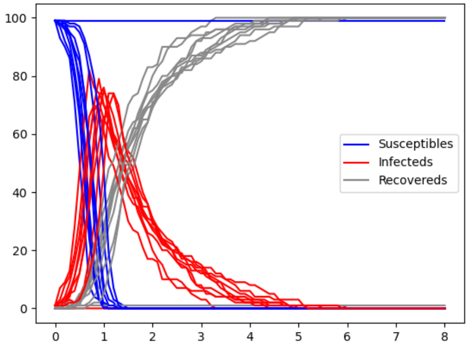
- Initial population: $S_0 = 99$, $I_0 = 1$, $R_0 = 0$
- Infection rate: $\beta = 0.1$ (which means $R_0 = 10$)
- Recovery rate: $\gamma = 1$

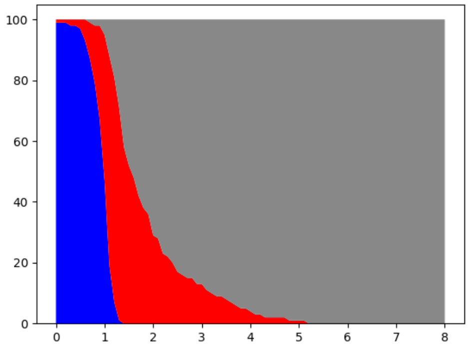
46.4 Compare the differences with the deterministic continuous model.

46.1 With the *discrete event* method, we recalculate the rates with every event, but with the τ-*leaping* method, we don't, we calculate how many events happen in a predetermined timestep size. Since in this model, the rate changes with each event, the discrete event method is more accurate.

46.2
$$\begin{aligned}\Pr(\text{new infectives}= k) &= \binom{S}{k}(\Delta t\beta I)^k(1-\Delta t\beta I)^{S-k}\\ \Pr(\text{new recovereds}= k) &= \binom{I}{k}(\Delta t\gamma)^k(1-\Delta t\gamma)^{I-k}\end{aligned}$$

46.3 The first graph is the result of 10 simulations with `SIR-stochastic-sol.ipynb`





46.4 Observe how on one of the simulations, no one else became infected, which would not be possible on the continuous model. Also note that the infected population actually becomes zero in finite time as opposed to asymptotically converging to 0 in the continuous model.

Monte Carlo Simulations and their accuracy.

The simulations we have done with the aquarium and SIR models are called Monte Carlo simulations. These rely on running the simulations again and again, which can take a long time. To compare the accuracy with other numerical methods:

- Monte Carlo has square-root convergence: To be 2x more accurate, we need to run it 4x more
- Euler has linear convergence: To be 2^1x more accurate, we need to run it 2x more
- Runge-Kutta 4 has 4th order convergence: To be 2^4x more accurate, we need to run it 2x more

They can be very convenient though to run through examples and try different strategies, e.g. for the aquarium store, changing the re-stocking policy required 1-2 different lines of coding.

Optimization with Stochastic Simulation.

We want to continue on the idea of Exercise 45.4.
We assume the following:

1. Each delivery has a cost $\$d$ independent of the number of aquariums shipped.

2. Each sale has a profit of $\$s$.

3. Each item in inventory has a small probability ρ of being damaged during the week and that would incur as a cost of $\$c$. The aquariums are damaged independently, so the number of damaged aquariums is binomial $K \sim B(X_n, \rho)$.

We want to optimize the profit.

48.1 What is the formula for the profit on week n with inventory X_n and demand D_n?

48.2 Let policy $p(x)$ be the number of aquariums that are reordered when there x aquariums in stock. What are all the possible values $p(x)$? How many possible policies are there?

48.3 If the store had a maximum of m aquariums in stock (instead of 3), how many different policies are there?

48.4 We can reduce these numbers, because the cost of ordering new aquariums is fixed and doesn't depend on how many aquariums are ordered.

So if we're willing to re-order to a certain stock L when the stock is x: $p(x) = ?$ then for $y < x$, we should have $p(y) = ?$

48.5 Based on the previous property, how many policies are there for a maximum of m aquariums?

48.1 We now have

$$R_n = \begin{cases} s \cdot \min\{D_n, X_n\} - d - c \cdot K_n & \text{if} \min\{D_n, X_n\} > 0 \\ -c \cdot K_n & \text{otherwise} \end{cases}$$

48.2 $p(3) = 0$, $p(2) \in \{0,1\}$, $p(1) \in \{0,1,2\}$, $p(0) \in \{0,1,2,3\}$, with a total number of possible policies of 4!.

48.3 If the maximum is m aquariums, then there are $(m+1)!$ possible policies, which quickly become too many to check.

48.4 So if we're willing to re-order to a certain stock L when the stock is x: $p(x) = L - x$ then for $y < x$, we should have $p(y) = L - y$?

48.5 So we have:

$$p(x) = \begin{cases} \max\{L - x, 0\} & \text{if } x < t \\ 0 & \text{if } x \geq t \end{cases}$$

where t is the level below which the store restocks and L is the level it tries to achieve.

The possible cases are: $t \in \{1, \ldots, m\}$ and $L \in \{t, \ldots, m\}$, so there are

$$\sum_{t=1}^{m} (m - t + 1) = \frac{1}{2} m(m+1).$$

Welford's Algorithm.
Suppose that the samples for the weekly profit are R_n for $n = 1, 2, \ldots$.
We compute two sequences μ_n and V_n to estimate the mean and variance of R.
The algorithm is given by:

$\mu_0 = V_0 = 0$;
forall $n \in \{1, 2, \ldots\}$ **do**
  $\delta_n = R_n - \mu_{n-1}$;
  $\mu_n = \mu_{n-1} + \frac{\delta_n}{n}$;
  $\delta'_n = R_n - \mu_n$;
  $V_n = V_{n-1} + \delta_n \delta'_n$;
end

Under some conditions on the random variable R:

- μ_n will converge to the mean of R;
- $\frac{V_n}{n}$ will converge to the variance of R.

We should ignore the first few weeks to make sure that we are sampling from the steady state distribution of profits.

We can now estimate the long term profit by simulating the Markov decision process for each policy.

- We can reduce the sampling error by simulating for a large number of weeks (instead of simulating several times).
- We can use Welford's algorithm to estimate the mean weekly profit and its variance.
- Even though Welford's algorithm reduces the computations needed to calculate the mean and variance, we still need to run the simulation for each possible policy and each week.

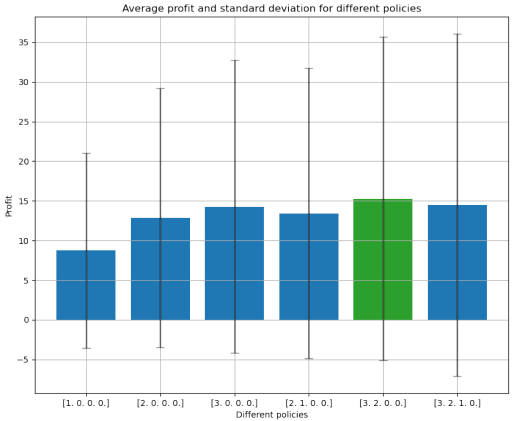
48.5 Use Welford's algorithm to calculate all mean and variance of the profit for all the possible policies with $m = 3$ (the original setting). Use the following values:

- $\lambda = 1$ expected customer per week
- Delivery cost $d = \$5$ and sale profit $s = \$20$
- Damage probability $\rho = 0.01$ and cost $c = \$60$

What is the best policy?

48.5 Running 1 000 000 weeks and ignoring the first 50 weeks over all 6 policies, we obtain the best policy:

- *Re-order if the stock* $X = 0$ *or* $X = 1$.
- The average weekly profit is \$ 15.28
- The standard deviation is \$ 20.38



Below is a graph of the average profits and their standard deviations for each re-stocking policy.

Code: `aquarium-MC-policies-sol.ipynb`

Q-learning Algorithm.

Let $Q(x,a)$ be the expected profit for the upcoming week with an inventory of x in the previous week and ordering a aquariums. We call this function the *quality of state x with action a*.

The optimal policy for state x is the action with the best quality: $p(x) = \text{argmax}_a Q(x,a)$.

We are going to determine $Q(s,a)$ using iterative refinement during a stochastic simulation of this Markov decision process.

---

for each week n **do**

$a_n = \text{random action}$

$$R_n = \underbrace{s\min(D_n, X_n)}_{\text{profit from sales}} - \underbrace{d(a_n > 0)}_{\text{delivery cost}} - \underbrace{cK_n}_{\text{damage cost}} \qquad \text{(reward/profit for the week)}$$

$$Q(X_n, a_n) \leftarrow (1-\alpha_n)Q(X_n, a_n) + \alpha_n \left(R_n + \gamma \max_a Q(X_{n+1}, a)\right)$$

end

The parameter α_n (*learning rate*) expresses the importance of the most recent observation relative to the prior knowledge represented in the current value of the quality function. We want $\alpha_n \to 0$ as the quality function converges.

The parameter γ (*discount factor*) expresses the notion that rewards later are worth less than rewards now. Since our process goes on forever, we want $\gamma < 1$ to endure convergence, otherwise the total rewards would grow indefinitely.

Let us now implement the Q-learning algorithm to calculate the optimal policies for different cases.

49.1 Use `aquarium-Qlearn.ipynb` to calculate the quality of state matrix Q for the original case $m = 3$, $\rho = 0.01$.

49.2 Use this to deduce the best policy in this case.

49.3 Use the same process to deduce the optimal policies for the cases $m \in \{1, \ldots 7\}$ and $\rho \in \{0.01, 0.03, 0.05\}$.

49.1 We obtain the matrix

$$Q = \begin{bmatrix} 83.98925114 & 91.15973818 & 95.97324719 & \mathbf{99.00385354} \\ 96.15340924 & 96.03440076 & \mathbf{98.52386915} & 0. \\ \mathbf{101.72981334} & 98.57852473 & 0. & 0. \\ \mathbf{104.05899297} & 0. & 0. & 0. \end{bmatrix}$$

49.2 To deduce the best policy:

- We need to find the value of a that maximizes the quality for each state
- We need to find the maximum for each row and identify the value of that column

For $m = 3$ and $\rho = 0.01$, we get:

- 3 2 0 0

which means that the store should order 3 aquariums when inventory is 0, 2 aquariums when inventory is 1, and not re-stock otherwise.

This matches our previous conclusions.

49.3 Here are my results:

m	policy for $\rho = 0.01$	policy for $\rho = 0.03$	policy for $\rho = 0.05$
1	1 0	1 0	1 0
2	2 1 0	2 0 0	2 0 0
3	3 2 0 0	3 2 0 0	2 0 0 0
4	4 3 0 0 0	3 2 0 0 0	2 0 0 0 0
5	5 4 0 0 0 0	3 2 0 0 0 0	2 0 0 0 0 0
6	4 4 0 0 0 0 0	3 2 0 0 0 0 0	2 0 0 0 0 0 0
7	5 4 0 0 0 0 0 0	3 2 0 0 0 0 0 0	2 0 0 0 0 0 0 0

Observe that because the Q-learning algorithm moves randomly from state to state and explores random actions a each iteration, the more possible states and actions, the *weeks* we need and the results are less conclusive.

After running the simulation a few times for $m = 6, 7$ and $\rho = 0.01$, the results varied between these policies:

- 4 4 0 $\cdots$
- 5 4 0 $\cdots$
- 6 4 0 $\cdots$