1.1 An introduction to optimization

Finding an optimal solution is ubiquitous in almost all areas ranging from science to engineering, economy, social studies, ...

For example,

- **Delivery route:** A package delivery man faces the problem of finding the optimal route of his daily deliveries to different addresses.
- **Log milling:** A woodmill worker always wants to cut a log in a way that optimizes the usage of each piece of log.
- The list can go on and on ......

Optimization processes can happen spontaneously and are observed almost everywhere in nature ...

For example,

- A drop of water free from other forces/influences (e.g. in space station) is spherical to minimize its surface tension.
- A closed system (i.e. with no energy or material exchange with outside world) always evolves toward a state that maximizes its entropy (a measure of the degree of disorder).
- In population genetics, natural selection tend to maximize the survival fitness of a species. (We shall demonstrate this in a simple mathematical model later on.)
- The list can go on and on ......
1.2 Review of optimization in calculus with simple examples

Eg.1 Given a fixed length of fence materials, try to encircle a fenced rectangular yard with the largest possible area. (Case 1: free from any other wall. Case 2: against an existing wall.)

Eg.2 Design the shape of a closed cylindrical container with a volume of 1 $m^3$ that minimizes the material needed to produce.

Now, let’s consider Eg.2 as a simple **mathematical modeling** problem and formulate the mathematics of it.

Let the radius be $R$ and height be $H$ (see figure).

**Q:** For what values of $R$ and $H$ its surface area is minimized?

Let’s formulate the quantity to be minimize: the surface area

$A = 2$ circular ends + 1 rectangular side.

Using simple geometric formula for areas of circle and rectangle, we obtain

$A = 2(\pi R^2) + 2\pi RH.$
• $A$ appears to depend on both $R$ and $H$ and is thus not a single variable function.

• But the constraint on the volume, $V = \pi R^2 H = 1$ makes the two depend on each other following the relation $H = \frac{1}{\pi R^2}$.

• Substitute this into the area express, we obtain a single variable function

$$A(R) = 2\pi R^2 + 2\pi R \frac{1}{\pi R^2} = 2\pi R^2 + \frac{2}{R}.$$  

which is now a single variable function!

Q: For what value of $R$ is $A$ minimum?

The question above now becomes a standard optimization problem in single variable calculus. Let’s refresh our memory with a brief review of optimization.
1.3 A brief review of optimization in single variable calculus.

Suppose \( f(x) \) is a smooth function (continuously differentiable to a certain degree) defined on interval \( I \).

**Absolute max (min):**
for some \( c \) in \( I \), if \( f(c) \geq (\leq) f(x) \) for all \( x \) in \( I \).

**Local max (min):**
for some \( c \) in \( I \), if \( f(c) \geq (\leq) f(x) \) for all \( x \) in \( I \) near \( c \).

When \( f(x) \) is smooth, as is often the case in real-world systems, local and absolute extrema happen either at a **critical point** (CP) or at an end point if \( I \) is a closed interval with finite ends.

**Critical point (CP):** for some \( c \) in \( I \), if \( f'(c) = 0 \), then it is a CP.

- A local extremum that is not an end point must be a CP.
- But the converse is not necessarily true.

Eg. \( f(x) = x^3 \), \( f'(x) = 3x^2 \) \( \Rightarrow f'(0) = 0 \)

but \( x = 0 \) is not a CP! (See figure).
**Brief summary:** for smooth functions defined on $I$

(1) Local extrema happen at CPs if located at the interior of $I$ (i.e. not an end point).

(2) If $I$ is closed, absolute and local extrema happen either at CPs or at end points.

(3) If $I$ is open, local extrema happen at CPs, absolute extrema either happen at local extrema or do not exist (for unbounded functions).

(4) If $I$ is half-closed, (2) or (3) can both apply.

(5) Solving an optimization problem typically involves finding the local extrema and compare the values of $f(x)$ there to the values at end points.
Back to Eg.2: The surface area

$$A = 2\pi R^2 + \frac{2}{R}$$

is defined on \((0, \infty)\) which is open!

Also, \(A \to \infty\) as \(R \to 0\) or \(\infty\). Thus, only absolute min can possibly exist which is what we look for.

$$A'(R) = 4\pi R - \frac{2}{R^2} = 0 \quad \Rightarrow \quad R_m = \left(\frac{1}{2\pi}\right)^{1/3} \approx 0.542 \ (m).$$

is the only CP. We can show that it is both the local and absolute min of \(A(R)\).

Either we show that \(A'(R) < 0\) for \(R < R_m\) and \(A'(R) > 0\) for \(R > R_m\), or we show that

$$A''(R_m) = 4\pi + \frac{4}{R_m^3} > 0 \quad \text{which is obviously true!}$$

Now, using the relation \(H = \frac{1}{\pi R^2}\), we find

$$H_m = \frac{1}{\pi R_m^2} = \frac{1}{\pi \left(\frac{1}{2\pi}\right)^{2/3}} = 2 \left(\frac{1}{2\pi}\right)^{1/3} = 2R_m \approx 1.084 \ (m).$$

Therefore, the shape that minimizes the surface area is the cylinder with its diameter equal to its height. It is easy to show this conclusion is independent of the volume of the container.
Eg.3 Maximized sales value. Weekly sales of a smart phone in one store increased from 250 to 275 when unit price is dropped from $500 to $450.

(a) Suppose that the price-demand relation follows a straight line, find its equation.

(b) Find the price that maximizes the weekly sales value.

Ans:

(a) Let \( n \) be the weekly sales number, and \( p \) be the unit price. Our goal is to find \( p = p(n) \). Since it is a straight line and we already know two points on the line \((n, p) = (250, 500)\) and \((275, 450)\), the slope is

\[
m = \frac{450 - 500}{275 - 250} = -2 \quad \Rightarrow \quad p(n) = 500 - 2(n - 250).
\]

A line with negative slope (decreasing) makes sense since larger sales number \( n \) is related to lower sales price \( p \).

(b) The weekly sales value that we need to maximize is

\[
V = np(n) = n(500 - 2(n - 250)) = 2n(500 - n)
\]

which is defined on a closed interval \([0, 500]\). Thus,

\[
V'(n) = 2(500 - n) + 2n(-1) = 1000 - 4n = 0 \quad \Rightarrow \quad n^* = 250
\]

is the only CP. It is a local max since \( V'' = -4 < 0 \) for all values of \( n \).

Since at the end points, \( V(0) = V(500) = 0 \),

\[
V(n^*) = 2(250)(500 - 250) = 125,000
\]
it must be the absolute max of $V(n)$. Substitute $n^* = 250$ into $p(n)$, we get the optimal price

$$p^* = 500 - 2(250 - 250) = 500.$$ 

The analysis shows that increased sales numbers at lower prices fail to generate higher sales value. $500 is actually the price that maximizes sales value.

Note that maximum sales profit is different since it would involve considering net profit margin of each unit sold. Maximizing sales profit would be a different optimization problem.
1.4 Typical steps in the construction of a model.

(1) Formulate the scientific question.

(2) Determine the basic ingredients (model variables, parameters, constraints, . . .).

(3) Describe the system being investigated schematically and/or verbally.

(4) Turn verbal/schematic model into quantitative equation(s).

(5) Analyze the equation(s) using numerical and/or analytical skills.

(6) Check the results against data and observations.

(7) Relate the results back to the biological question.
1.5 Application to modeling evolution & population genetics

“As many more individuals of each species are born than can possibly survive, and as, consequently, there is a frequently recurring struggle for existence, it follows that any being, if it vary slightly in any manner profitable to itself, under the complex and sometimes varying conditions of life, will have a better chance of surviving, and thus be naturally selected.”

– Charles Darwin, The Origin of Species (1859)
Key characteristics of evolution

- Organisms vary in their traits.
- Not all individuals survive and certain traits improve fitness.
- Traits may be inherited.

**Fitness:** can here be technically and narrowly defined as the average number of adult/reproductive descendants reproduced per individual from one generation to the next. (It is often referred to as the *basic reproduction number* or *basic reproductive rate* in areas like epidemiology.)

To put it in mathematical terms: let $N_a(t)$ and $N_a(t + 1)$ be the numbers of adult individuals with trait $a$ in $t^{th}$ and $t + 1^{th}$ generations respectively, then the average fitness of individuals with trait $a$ is defined as

$$
\bar{F}_a = \frac{N_a(t + 1)}{N_a(t)}.
$$

Parents with characteristics that improve fitness are likely to have more offsprings. These characteristics therefore increase in frequency over time leading the population to evolve in favour of individuals with those traits.
Important questions in evolution

- Why is polymorphism maintained?
- What forces are responsible for the maintenance of polymorphism?
- What exact roles different mechanisms (e.g. selection, drift, and mutation) play for the maintenance of polymorphism?

**Selection:** difference in survivability in different genotypes.

**Drift:** changes in frequency of different genotypes caused by random sampling effects.

**Mutation:** random occurrence of new genes.
Questions of interest in evolution and ecology

- Under what conditions will natural selection produce traits that are optimal?
- Why are all organisms within a population not of the optimal type: why is variability maintained?
- How fast is the response to selection? How quickly does the population evolve toward the optimum?
A model of evolution in absence of selection

Consider one-locus of a population of diploid individuals with two variant alleles $A$ and $a$.

- **Gene**: Segment of the DNA, generally a region that codes for a single protein.
- **Locus**: A site on a chromosome (usually synonymous with gene).
- **Allele**: A variant of a gene (a particular sequence).
- **Haploid/Diploid**: individuals that carry one copy/two copies of each gene.
- **Gamete**: The reproductive cell of a diploid sexual organism (e.g. sperm or egg).
- **Genotype**: The alleles carried by an individual at a gene.
- **Homozygote/Heterozygote**: Individual that carries two identical/different alleles.
Step 1: Formulate the scientific question

How do gene frequencies change over time in the absence of natural selection?
Step 2: Determine the basic ingredients

Model variables
- \( x = \) frequency of \( AA \) individuals.
- \( y = \) frequency of \( Aa \) individuals.
- \( z = \) frequency of \( aa \) individuals.

Constraints on model variables
- \( 0 \leq x, y, z \leq 1. \)
- \( x + y + z = 1. \)

Model variable(s) of crucial importance. Let the size of the population be \( N \).
- Frequency of \( A \) allele:
  \[ p = \frac{\text{total number of } A \text{ alleles}}{\text{total number of alleles}} = \frac{2Nx + Ny}{2N} = x + \frac{y}{2}. \]
- Frequency of \( a \) allele:
  \[ q = \frac{\text{total number of } a \text{ alleles}}{\text{total number of alleles}} = \frac{2Nz + Ny}{2N} = z + \frac{y}{2}. \]
- \( p + q = 1 \) or \( q = 1 - p. \)
- We assume that mating is completely random.
Step 4: Time evolution of model variables

At $t = 0$: model variables are $x_0$, $y_0$, $z_0$, thus

$$p_0 = x_0 + \frac{y_0}{2}, \quad q_0 = z_0 + \frac{y_0}{2}.$$ 

At $t = 1$: based on the mating table

- $x_1 = p_0^2 = (x_0 + \frac{y_0}{2})^2$;
- $y_1 = 2p_0q_0 = 2(x_0 + \frac{y_0}{2})(z_0 + \frac{y_0}{2})$;
- $z_1 = q_0^2 = (z_0 + \frac{y_0}{2})^2$.

Therefore,

- $p_1 = x_1 + \frac{y_1}{2} = p_0^2 + p_0q_0 = p_0(p_0 + q_0) = p_0$;
- $q_1 = z_1 + \frac{y_1}{2} = q_0^2 + p_0q_0 = q_0(q_0 + p_0) = q_0$. 

Step 5: Analyze the model equation(s)

\[ p_{t+1} = p_t = p_0, \quad (q_{t+1} = q_t = q_0). \]

And that

- \( x_{t+1} = p_t^2 = p_0^2; \)
- \( y_{t+1} = 2p_tq_t = 2p_0q_0; \)
- \( z_{t+1} = q_t^2 = q_0^2. \)
Step 6: Check the results against data

Is there any data that support unchanged allele frequencies?

*Answer:* Yes!

Therefore,

- \( x_0 = 0.292; \)
- \( y_0 = 0.496; \)
- \( z_0 = 0.212. \)
- \( p_0 = x_0 + y_0/2 = 0.540; \)
- \( q_0 = 1 - p_0 = 0.460. \)
- \( x_1 = p_0^2 = 0.2916; \)
- \( y_1 = 2p_0q_0 = 0.4968; \)
- \( z_1 = q_0^2 = 0.2116. \)

Observation agrees with unchanged allele frequency, indicating that an equilibrium (called Hardy-Weinberg equilibrium) has been reached.
Step 7: Relate back to the biological question

Hardy-Weinberg Equilibrium: A population is at the Hardy-Weinberg equilibrium if the allele frequencies do not change after one single generation.

Hardy-Weinberg Law: Random mating in the absence of selection leads to unchanged allele frequencies. Therefore, the Hardy-Weinberg equilibrium is reached in one single step.

From a mathematical point of view

• The discrete model without selection is described by a single-variable, linear map:

\[
p_{t+1} = p_t
\]

• All other variables can be derived from this variable:

\[
\begin{align*}
q_t &= 1 - p_t; \\
x_t &= p_t^2; \\
y_t &= 2p_tq_t; \\
z_t &= q_t^2.
\end{align*}
\]
Assume that

- $N(t)=$ total number of adult individuals at time $t$.
- $p, q =$ frequencies of alleles $A, a$ in adults at time $t$
- $n_{AA}(t + 1) = p^2 N(t)$
  = number of AA juveniles at time $t + 1$.
- $n_{Aa}(t + 1) = 2pq N(t)$
  = number of Aa juveniles at time $t + 1$.
- $n_{aa}(t + 1) = q^2 N(t)$
  = number of aa juveniles at time $t + 1$. 
One-Locus Diploid Selection Model

Introducing the survival probabilities

- \( V_{AA} \) = survival probability of AA juveniles to adulthood.
- \( V_{Aa} \) = survival probability of Aa juveniles to adulthood.
- \( V_{aa} \) = survival probability of aa juveniles to adulthood.

Therefore,

- \( N_{AA}(t + 1) = n_{AA}V_{AA} = p^2N(t)V_{AA} \)
  = number of adult AA individuals at time \( t + 1 \).
- \( N_{Aa}(t + 1) = n_{Aa}V_{Aa} = 2pqN(t)V_{Aa} \)
  = number of adult Aa individuals at time \( t + 1 \).
- \( N_{aa}(t + 1) = n_{aa}V_{aa} = q^2N(t)V_{aa} \)
  = number of adult aa individuals at time \( t + 1 \).
Time evolution of allele frequencies

\[ p(t + 1) = \frac{2N_{AA}(t + 1) + N_{Aa}(t + 1)}{2[N_{AA}(t + 1) + N_{Aa}(t + 1) + N_{aa}(t + 1)]} \]

Denote \( p(t) \) and \( p(t + 1) \) by \( p \) and \( p' \) respectively,

\[ p' = \frac{2p^2N_tV_{AA} + 2pqN_tV_{Aa}}{2[p^2N_tV_{AA} + 2pqN_tV_{Aa} + q^2N_tV_{aa}]} \Rightarrow \]

\[ p' = \frac{p^2V_{AA} + pqV_{Aa}}{p^2V_{AA} + 2pqV_{Aa} + q^2V_{aa}} = \frac{p\bar{V}_A}{\bar{V}}. \]

- \( q = 1 - p \) and \( q' = 1 - p' \).
- This is a nonlinear map/difference equation.
- Fitness: \( \bar{V} = p^2V_{AA} + 2pqV_{Aa} + q^2V_{aa} \).
- Fitness of allele A: \( \bar{V}_A = pV_{AA} + qV_{Aa} \).
- Often, no analytical solution is achievable.
Notice that

- The time evolution equations we obtained in the last two models are both **discrete-time dynamical systems**.
- Such equations are sometimes referred to as maps, recurrence equations, iterations, or **difference equation**.
- In the first model, the equation is **linear** but the second equation is highly **nonlinear**.
- Linear difference equations can readily be solved analytically. However, the nonlinear ones are often difficult or even impossible to solve analytically in closed forms.
- Apparently simple looking nonlinear maps can generate really complicated phenomena such as deterministic stochasticity and chaos.

- **Dynamical system**: can be generally understood as a system in which major variable(s) vary as a function of time in the state space.
- **Discrete time**: when time varies in finite steps but not in a continuous way.
- **Difference equation**: an equation that relates the value of a discrete time variable at time $n$ to its value at one or several previous times. _Difference equations relate to differential equations as discrete time models relate to continuous time models._
- **Linear**: an equation in which the dependent variable (which is often the unknown to be solved) appears only as multiples of itself or its derivatives.
- **Nonlinear**: an equation that is not linear.
We shall introduce some basic concepts and skills in the analysis of discrete time difference equations (maps) in the next lecture and then use the skills to analyze the “one-locus diploid selection model” we constructed in this lecture.