Lecture 9: introduction to nonlinear optimization

- Convex functions
- Convex sets
- Reminder: derivative
- Gradient descent
- Problems of gradient descent

Convex function
Convex function

\[ f(x) = f(\lambda a + (1 - \lambda)b) \leq \lambda f(a) + (1 - \lambda)f(b) \]

Non convex function

\[ f(x) = f(\lambda a + (1 - \lambda)b) \leq \lambda f(a) + (1 - \lambda)f(b) \]
Concave function

\[ f(x) = f(\lambda a + (1-\lambda)b) \geq \lambda f(a) + (1-\lambda)f(b) \quad 0 \leq \lambda \leq 1 \]

Definitions / facts, common mistakes

Definitions/facts
- If \( f \) is convex, \(-f\) is concave
- If \( f \) is concave, \(-f\) is convex

Common mistakes
- A nonconvex function is not necessarily concave
- A nonconcave function is not necessarily convex
- A function can be neither concave nor convex
- A function can be concave and convex
Functions not defined for all $x$

Functions not defined for all $x$ (convex)
Functions not defined for all x (convex)

\[ f(x) = +\infty \]
Functions not defined for all $x$ (not convex)

Convex sets

Definition:
Convex set: for all $A$ and $B$ in the set, if $A$ and $B$ are in the set, $\lambda A + (1-\lambda)B$ is also in this set, for $0 \leq \lambda \leq 1$
Convex sets

Definition:
Convex set: for all A and B in the set, if A and B are in the set, $\lambda A + (1 - \lambda)B$ is also in this set, for $0 \leq \lambda \leq 1$

This set is not convex

Famous convex sets

- Soccer ball
- Rugby ball
- The Great Pyramid of Giza
- The Pentagon
Famous non convex sets

Why do we use ‘convex’ for functions and sets

The epigraph (i.e. points above the graph) of a convex function is a convex set.

Convex function

Non convex function
Why do we use ‘convex’ for functions and sets

The epigraph (i.e. points above the graph) of a convex function is a convex set.

Convex sets and functions: basic properties

Non convex function

Convex function

Local minimum might not be a global minimum

Local minimum is a global minimum
Convex sets and functions: basic properties

Convex function

Minimum might not be unique

Local minimum is a global minimum

Convex optimization programs

Convex function

\[
\begin{align*}
\text{min:} & \quad f(x) \\
\text{s.t.:} & \quad g(x) \leq 0
\end{align*}
\]

f and g are convex functions, defined on convex sets

Convex optimization programs are “easy” problems, compared to general optimization programs

Local minimum is a global minimum
The derivative of a function at a point represents its slope.
Gradient descent (conceptual description)

You want to find the minimum of a function, starting from a guess, assuming that you cannot depict the graph of the function, for example the following function

\[ f(x) = \exp(\sin(x^2)) + \sqrt{x^4 + 3 \sin(\exp(-\frac{1}{1+\epsilon|x|}))} \]

Idea:
1) Make a guess
2) Compute the derivative at this point (i.e. the slope)
3) Follow the direction of the slope (i.e. descend)
4) Stop when the slope is zero, i.e. it does not go downhill

Gradient descent (illustration)
Gradient descent (illustration)

Gradient descent (illustration)
Gradient descent (illustration)

1. **Starting Point**: Choose an initial guess for the minimum ($m_0$).
2. **Gradient Calculation**: Compute the gradient of the function $f(x)$ at the current point (e.g., $f(m)$).
3. **Direction**: Use the gradient to determine the direction of the next step (e.g., $f(x)$).
4. **Update**: Move in the negative gradient direction to find the next point (e.g., $m$).
5. **Convergence**: Repeat the process until the function value $f(m)$ is sufficiently small or stops changing significantly (e.g., $f(m)$.

Graphically, this process is depicted with a graph showing the function $f(x)$, the gradient descent path, and the iterative updates towards the minimum.
Problem 1: non convex function

\[ f(m) \]

"good" guess

"bad" guess
Problem 2: how to stop?

What if you cannot figure out where the stop point is?

Problem 3: how to hit the wall?

"algebraic minimum"
Problem 3: how to hit the wall?

“actual minimum”