Plan for today

• Nonlinear Optimization
  • Unconstrained
    • Verifying convexity
    • Minimizing Convex Functions
  • Constrained
    • Equality constraints: Lagrangian Method
    • Inequality constraints: KKT Conditions Method
Recap

NONLINEAR OPTIMIZATION

... where we see nonlinear optimization models and define convex functions

Nonlinear Optimization = Non Linear Programming (NLP)
Linear vs Non-Linear functions

- A function $f(x)$ is **linear** if it satisfies proportionality and additivity.

  $$f(x) = 7x - 4$$
  $$f(x) = 3x_1 + 7x_2 - 15x_3$$

- A function $f(x)$ is **non-linear** if it violates proportionality or additivity.

  $$f(x) = x^2 - x + 4$$
  $$f(x) = x_1^2 - 16x_1^3 + x_1x_2$$
NLP: UNCONSTRAINED

Basic Concepts in NLP
Local vs Global Minima

- **Local Minimum:**
  - A point is a local minimum if there are no better (feasible) solutions “nearby”
  - Formally: A point $x^*$ is a local minimum if there exists some $\epsilon > 0$ such that $f(x^*) \leq f(x)$ for all feasible $x$ such that $||x - x^*|| < \epsilon$

- **Global Minimum:**
  - A point is a global minimum if there are no better (feasible) solutions anywhere
  - Formally: A point $x^*$ is a global minimum if $f(x^*) \leq f(x)$ for all feasible $x$
How to find stationary points for multivar functions?

• Suppose that our mathematical program is of the form
  \[
  \min f(x) \\
  x \text{ unrestricted}
  \]
  where \( f(x) \) is a multi var differentiable function

• From calculus, we can find stationary points by solving for \( x \) from
  \[
  \nabla f(x) = 0
  \]
  where the gradient of \( f \) is defined as \( \nabla f(x) := \left( \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n} \right) \)

• Checking if a stationary point is a local min/max is much harder for multivar functions
Local vs Global Minima/Maxima

- **Local Minimum**:
  - A point is a local minimum if there are no better (feasible) solutions “nearby”
  - Formally: A point $x^*$ is a local minimum if there exists some $\epsilon > 0$ such that $f(x^*) \leq f(x)$ for all feasible $x$ such that $\|x - x^*\| < \epsilon$

- **Global Minimum**:
  - A point is a global minimum if there are no better (feasible) solutions anywhere
  - Formally: A point $x^*$ is a global minimum if $f(x^*) \leq f(x)$ for all feasible $x$

**Question**: Suppose we know that $x^*$ is a stationary point. Can we tell whether it is a global minimum? Convexity helps!

- If $f(x)$ is a convex function, then any stationary point is a global minimum
- If $f(x)$ is a concave function, then any stationary point is a global maximum
Convex Functions

A function $f(x)$ is a

1. **convex function** if
   for each pair of points $a$ and $b$ and every $0 < \lambda < 1$:
   
   $$f(\lambda a + (1 - \lambda)b) \leq \lambda f(a) + (1 - \lambda)f(b)$$

2. **strict convex function** if
   for each pair of points $a$ and $b$ and every $0 < \lambda < 1$:
   
   $$f(\lambda a + (1 - \lambda)b) < \lambda f(a) + (1 - \lambda)f(b)$$

3. **concave function** if
   for each pair of points $a$ and $b$ and every $0 < \lambda < 1$:
   
   $$f(\lambda a + (1 - \lambda)b) \geq \lambda f(a) + (1 - \lambda)f(b)$$

4. **strict concave function** if
   for each pair of points $a$ and $b$ and every $0 < \lambda < 1$:
   
   $$f(\lambda a + (1 - \lambda)b) > \lambda f(a) + (1 - \lambda)f(b)$$
Local vs Global Minima/Maxima

- **Local Minimum:**
  - A point is a local minimum if there are no better (feasible) solutions “nearby”
  - Formally: A point $x^*$ is a local minimum if there exists some $\epsilon > 0$ such that $f(x^*) \leq f(x)$ for all feasible $x$ such that $\|x - x^*\| < \epsilon$

- **Global Minimum:**
  - A point is a global minimum if there are no better (feasible) solutions anywhere
  - Formally: A point $x^*$ is a global minimum if $f(x^*) \leq f(x)$ for all feasible $x$

**Question:** Suppose we know that $x^*$ is a stationary point. Can we tell whether it is a global minimum?  
Convexity helps!

- If $f(x)$ is a convex function, then any stationary point is a global minimum
- If $f(x)$ is a concave function, then any stationary point is a global maximum
HOW TO VERIFY CONVEXITY?

- By definition (often hard)
- Derivative tests
How to verify convexity for \textbf{multi var} functions?

- For a single variable twice differentiable function $f(x)$
  - If $f''(x) \geq 0$ for all $x$, then $f(x)$ is a convex function
  - If $f''(x) > 0$ for all $x$, then $f(x)$ is a strictly convex function
  - If $f''(x) \leq 0$ for all $x$, then $f(x)$ is a concave function
  - If $f''(x) < 0$ for all $x$, then $f(x)$ is a strictly concave function

- A multi variable twice-differentiable function $f(x)$ is \textbf{convex} if all \textbf{eigenvalues} of the \textbf{Hessian matrix} $\nabla^2 f(x)$ are non-negative
  - The $i$-th row, $j$-th col entry of the \textbf{Hessian matrix} is obtained by
    \[
    [\nabla^2 f(x)]_{i,j} := \frac{\partial^2 f}{\partial x_i \partial x_j}
    \]
  - Recall: Eigenvalues of a matrix $M$ are the roots of the equation $\det(M - yI) = 0$
PUTTING THE CONCEPTS TOGETHER

Unconstrained Minimization
Minimizing Single Variable Twice-Differentiable Functions

\[
\begin{array}{|c|}
\hline
\min f(x) \\
x \text{ unrestricted} \\
\hline
\end{array}
\]

1. Find a stationary point \( x^* \)
   \( \ldots \) by solving \( f'(x^*) = 0 \)

2. Verify that \( f(x) \) is convex
   \( \ldots \) by verifying if \( f''(x) > 0 \) for all \( x \)
Minimizing Single Variable Twice-Differentiable Functions

$$\min f(x) = x^2$$
$$x \text{ unrestricted}$$

1. Find a stationary point $x^*$
   ... by solving $f'(x^*) = 0$

$$f'(x) = 2x$$
$$x^* = 0$$

2. Verify that $f(x)$ is convex
   ... by verifying if $f''(x) > 0$ for all $x$

... Have done this already for this example
Minimizing Multi Variable Twice-Differentiable Functions

\[
\begin{align*}
\min f(x) \\
x = (x_1, x_2, \ldots, x_n) \text{ unrestricted}
\end{align*}
\]

1. Find a point \( x^* \) where the gradient is zero
   … by solving \( \nabla f(x) = 0 \) where, the gradient of \( f \) is defined as \( \nabla f(x) := \left( \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n} \right) \)

2. Verify that \( f(x) \) is convex
   … by verifying if all eigenvalues of the **Hessian matrix** \( \nabla^2 f(x) \) are non-negative
Minimizing Multi Variable Twice-Differentiable Functions

\[
\begin{align*}
\min f(x_1, x_2) &= x_1^2 + x_2^2 \\
x &= (x_1, x_2) \text{ unrestricted}
\end{align*}
\]

1. Find a point \(x^*\) where the gradient is zero
   ... by solving \(\nabla f(x) = 0\) where,
   
   the gradient of \(f\) is defined as \(\nabla f(x) := \left( \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n} \right)\)

   \[
   \nabla f(x) = \left( \frac{\partial f}{\partial x_1} = 2x_1, \frac{\partial f}{\partial x_2} = 2x_2 \right)
   \]
   
   \[x^* = (x_1^* = 0, x_2^* = 0)\]

2. Verify that \(f(x)\) is convex
   ... by verifying if all eigenvalues of the **Hessian matrix** \(\nabla^2 f(x)\) are non-negative

   ... Have done this already for this example
Minimizing Single Variable Twice-Differentiable Functions

\[
\min_{x \text{ unrestricted}} f(x)
\]

1. Find a stationary point \( x^* \)
   \[ \quad \text{... by solving } f'(x^*) = 0 \]

2. Verify that \( f(x) \) is convex
   \[ \quad \text{... by verifying if } f''(x) > 0 \text{ for all } x \]

If \( f(x) \) is not convex, then
1. Find all stationary points and identify local minimizers
2. Evaluate the function at all local minimizers to identify the global minimizer
Minimizing Multi Variable Twice-Differentiable Functions

\[
\begin{align*}
\min f(x) \\
\mathbf{x} = (x_1, x_2, \ldots, x_n) \text{ unrestricted}
\end{align*}
\]

1. Find a point \( x^* \) where the gradient is zero
   … by solving \( \nabla f(x) = 0 \) where,
   the gradient of \( f \) is defined as \( \nabla f(x) := \left( \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n} \right) \)

2. Verify that \( f(x) \) is convex
   … by verifying if all eigenvalues of the \textbf{Hessian matrix} \( \nabla^2 f(x) \) are non-negative

\[\text{If } f(x) \text{ is not convex, then problem is very difficult}\]

Machine Learning Applications!
Maximizing vs Minimizing
Convex vs Concave

- Maximizing $f(x)$ is equivalent to minimizing $-f(x)$

- Also note that $f(x)$ is concave if and only if $-f(x)$ is convex

- So, concave maximization is equivalent to convex minimization
NLP: CONSTRAINED

... where we see optimality conditions for multivariate constrained nonlinear optimization problems
NLP: CONSTRAINED

With only equality constraints

- The Lagrangian Method
Constrained Optimization with Equality Constraints

• Suppose we have an optimization problem of the following type:
  \[
  \max f(\mathbf{x}) \\
  g_i(\mathbf{x}) = b_i \text{ for } i = 1, \ldots, m
  \]
  where \( f(\mathbf{x}) \) and any of the \( g_i(\mathbf{x}) \) may be non-linear and \( \mathbf{x} = (x_1, x_2, \ldots, x_n) \)

• Given this problem, we can write a related **unconstrained** optimization problem:

  \[
  \max_{\mathbf{x}, \lambda} h(\mathbf{x}, \lambda) \\
  h(\mathbf{x}, \lambda) := f(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i (g_i(\mathbf{x}) - b_i)
  \]

  where we have added \( m \) new variables \( \lambda_1, \lambda_2, \ldots, \lambda_m \) called **Lagrange multipliers**

• The objective function \( h(\mathbf{x}, \lambda) \) is called the **Lagrange function**
Constrained Optimization with Equality Constraints

\[
\begin{align*}
\max & \quad f(x) \\
\text{s.t.} & \quad g_i(x) = b_i \quad \text{for } i = 1, \ldots, m
\end{align*}
\]

We can find the stationary points of \( h(x, \lambda) \) in the same way that we found the stationary points of unconstrained multi-var NLP

- We now have \( n + m \) decision variables in \( h(x, \lambda) \)
- So take the gradient of \( h \), set it to zero, and solve for \( x \) and \( \lambda \)
- Suppose \( (x^*, \lambda^*) = (x_1^*, x_2^*, \ldots, x_n^*, \lambda_1^*, \lambda_2^*, \ldots, \lambda_m^*) \) is a stationary point
- Recall that \( \frac{\partial h(x, \lambda)}{\partial \lambda_i} = g_i(x) - b_i = 0 \) at any stationary point of \( h(x, \lambda) \)
- So every stationary point corresponds to a feasible solution to our constrained optimization problem

\[
\begin{align*}
\max_{x, \lambda} & \quad h(x, \lambda) \\
\text{s.t.} & \quad g_i(x) = b_i \quad \text{for } i = 1, \ldots, m
\end{align*}
\]

\[
h(x, \lambda) := f(x) - \sum_{i=1}^{m} \lambda_i (g_i(x) - b_i)
\]

The stationary points of the Lagrangian are known as the critical points of the constrained optimization problem

Note: If \( x^* = (x_1^*, x_2^*, \ldots, x_n^*) \) is a constrained optimum, then there must be some \( \lambda^* = (\lambda_1^*, \lambda_2^*, \ldots, \lambda_m^*) \) such that \( (x^*, \lambda^*) \) is a stationary point of the Lagrangian
  - i.e., we know that one of the stationary points should be an optimum for the constrained problem

After finding the stationary points of the Lagrangian, pick the maximizer among them
  - Caution: Remember to verify if the picked point is a maximizer/minimizer
Consider the following constrained NLP:

\[
\begin{align*}
    \max & \quad x_1 + x_2 \\
    \text{s.t.} & \quad x_1^2 + x_2^2 = 1
\end{align*}
\]

The Lagrangian function is

\[
h(x, \lambda) = x_1 + x_2 - \lambda_1 (x_1^2 + x_2^2 - 1)
\]

Find stationary points by setting \( \nabla h(x, \lambda) = 0 \):

\[
\begin{align*}
    \frac{\partial h(x, \lambda)}{\partial x_1} &= 1 - 2\lambda_1 x_1 = 0 \\
    \frac{\partial h(x, \lambda)}{\partial x_2} &= 1 - 2\lambda_1 x_2 = 0 \\
    \frac{\partial h(x, \lambda)}{\partial \lambda_1} &= -(x_1^2 + x_2^2 - 1) = 0
\end{align*}
\]

We get a system of nonlinear equations, which may be difficult to solve
Constrained Optimization with Equality Constraints

• From the first two equations, we have:
  \[
  \frac{\partial h(x, \lambda)}{\partial x_1} = 1 - 2\lambda_1 x_1 = 0, \text{ implies } x_1^* = \frac{1}{2\lambda_1^*}
  \]
  \[
  \frac{\partial h(x, \lambda)}{\partial x_2} = 1 - 2\lambda_1 x_2 = 0, \text{ implies } x_2^* = \frac{1}{2\lambda_1^*}
  \]

• Substituting into the third equation, we have:
  \[
  \frac{\partial h(x, \lambda)}{\partial \lambda_1} = -(x_1^*^2 + x_2^*^2 - 1) = \frac{1}{(2\lambda_1^*)^2} + \frac{1}{(2\lambda_1^*)^2} - 1 = 0, \text{ implies } \lambda_1^* = \pm \frac{1}{\sqrt{2}}
  \]

• Taking each of the two possible values for \( \lambda_1^* \), we find two stationary points
  • \((x^*, \lambda^*) = \left( \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right) \), i.e., \((x_1^*, x_2^*) = \left( \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right)\)
  • \((x^*, \lambda^*) = \left( -\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right) \), i.e., \((x_1^*, x_2^*) = \left( -\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right)\)

• The stationary point with the best objective value is \((x_1^*, x_2^*) = \)

Is this a maximizer or a minimizer?
Constrained Optimization with Equality Constraints

- Feasible region:

\[
\begin{align*}
\max x_1 + x_2 \\
x_1^2 + x_2^2 &= 1
\end{align*}
\]
Constrained Optimization with Equality Constraints

- Critical points:

Example: Verification by the Graphical Method

\[
\begin{align*}
\max x_1 + x_2 \\
x_1^2 + x_2^2 &= 1
\end{align*}
\]
Constrained Optimization with Equality Constraints

- Iso-line:

\[ x_1 + x_2 = 0 \]

Solving graphically, we observe that

- \( \left( \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right) \) is a global maximum
- \( \left( -\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right) \) is a global minimum

Example: Verification by the Graphical Method

\[
\max x_1 + x_2 \\
x_1^2 + x_2^2 = 1
\]
NLP: CONSTRAINED

With **inequality** constraints

- The KKT Conditions Method
Multi-var Constrained Opt

- Suppose we want to solve
  \[
  \max f(x) \quad \text{subject to} \quad g_i(x) \leq b_i \quad \text{for } i = 1, \ldots, m
  \]
  where \( f(x) \) and any of the \( g_i(x) \) may be non-linear
  and \( x = (x_1, x_2, \ldots, x_n) \)

We will talk only about constrained optimization problems of the above form for the sake of following a convention.
• Suppose we want to solve
\[
\max f(x) \text{ subject to } g_i(x) \leq b_i \text{ for } i = 1, \ldots, m
\]
where \( f(x) \) and any of the \( g_i(x) \) may be non-linear and \( x = (x_1, x_2, \ldots, x_n) \)

We will talk only about constrained optimization problems of the above form for the sake of following a convention.
KKT conditions

Suppose we have an optimization problem of the following type:

$$\max f(x)$$

$$g_i(x) \leq b_i \text{ for } i = 1, \ldots, m$$

where $f(x)$ and any of the $g_i(x)$ may be non-linear and $x = (x_1, x_2, \ldots, x_n)$

**Theorem:** If $x = (x_1, x_2, \ldots, x_n)$ is a local or global optimum of the constrained problem, then there must be values $u = (u_1, u_2, \ldots, u_m)$ such that:

1. $$\frac{\partial f(x)}{\partial x_j} - \sum_{i=1}^{m} u_i \frac{\partial g_i(x)}{\partial x_j} = 0 \text{ for } j = 1,2,\ldots,n$$
2. $$g_i(x) \leq b_i \text{ for } i = 1,2,\ldots,m$$
3. $$u_i \geq 0 \text{ for } i = 1,2,\ldots,m$$
4. $$u_i(g_i(x) - b_i) = 0 \text{ for } i = 1,2,\ldots,m$$

Conditions 1, 2, 3, and 4 are called **Karush-Kuhn-Tucker (KKT) conditions**

**NOTE:** Theorem also requires $f$ and $g$ to satisfy some “regularity conditions”, which will be satisfied by the functions encountered within the scope of this course
Constrained Optimization with Inequality Constraints (KKT conditions)

• Like the Lagrangian, we can use these conditions to identify critical points that could be local or global optima

• These points are called **KKT points**

• After computing the KKT points, pick the maximizer among them
  • Caution: Remember to verify if the picked point is a maximizer/minimizer
  • How to verify?
    1. Graphical method
    2. If the objection function $f$ is concave and all constraint functions $g_i$ are convex, then the KKT point that you found is a maximizer
Constrained Optimization with Inequality Constraints (KKT conditions)

- Consider the following constrained NLP:

\[
\begin{align*}
&\text{max } x_1 + x_2 \\
&x_1^2 + x_2^2 \leq 1 \\
&x_1 \leq \frac{1}{2}
\end{align*}
\]

\[
\begin{align*}
f(x) &= x_1 + x_2 \\
g_1(x) &= x_1^2 + x_2^2, \quad b_1 = 1 \\
g_2(x) &= x_1, \quad b_2 = \frac{1}{2}
\end{align*}
\]

- The KKT conditions for this problem are:

1. \[
\frac{\partial f(x)}{\partial x_1} - u_1 \frac{\partial g_1(x)}{\partial x_1} - u_2 \frac{\partial g_2(x)}{\partial x_1} = 0
\]

2. \[
\frac{\partial f(x)}{\partial x_2} - u_1 \frac{\partial g_1(x)}{\partial x_2} - u_2 \frac{\partial g_2(x)}{\partial x_2} = 0
\]

3. \[
g_1(x) \leq b_1
\]

4. \[
g_2(x) \leq b_2
\]

5. \[
u_1 (g_1(x) - b_1) = 0
\]

6. \[
u_2 (g_2(x) - b_2) = 0
\]

7. \[u_1, u_2 \geq 0\]

- Solving these equations provides the KKT points
Constrained Optimization with Inequality Constraints (KKT conditions)

- To solve, try different cases for the $u$ variables
- First choice: $u_1 = 0 \rightarrow$ Impossible (violates equation 2)
- What if $u_1 > 0$ and $u_2 = 0$
  - $x_1 = \frac{1}{2u_1}$ (by eqn 1)
  - $x_2 = \frac{1}{2u_1}$ (by eqn 2)
  - $u_1 = \frac{1}{\sqrt{2}}$ (by eqn 5)
  - Then $x_1 = \frac{1}{2u_1} = \frac{1}{\sqrt{2}}$ which contradicts constraint 4
- What if $u_1 > 0$ and $u_2 > 0$?
  - $x_1 = \frac{1}{2}$ (by eqn 6)
  - $x_2 = \pm \frac{\sqrt{3}}{2}$ (by eqn 5 with $u_1 > 0$)
  - If $x_2 = -\frac{\sqrt{3}}{2}$, $u_1 = -\frac{1}{\sqrt{3}}$ (by eqn 2), which contradicts $u_1 > 0$
  - If $x_2 = \frac{\sqrt{3}}{2}$, $u_1 = \frac{1}{\sqrt{3}}$ (by eqn 2)
  - $u_2 = 1 - \frac{1}{\sqrt{3}}$ (by eqn 1)
- KKT point is $(x_1, x_2, u_1, u_2) = \left(\frac{1}{2}, \frac{\sqrt{3}}{2}, \frac{1}{\sqrt{3}}, 1 - \frac{1}{\sqrt{3}}\right)$

1. $1 - 2u_1x_1 - u_2 = 0$
2. $1 - 2u_1x_2 = 0$
3. $x_1^2 + x_2^2 \leq 1$
4. $x_1 \leq \frac{1}{2}$
5. $u_1(x_1^2 + x_2^2 - 1) = 0$
6. $u_2\left(x_1 - \frac{1}{2}\right) = 0$
7. $u_1, u_2 \geq 0$

Note: This is the unique KKT point. Need to verify if it is a maximizer or a minimizer.
Constrained Optimization with Inequality Constraints (KKT conditions)

• Like the Lagrangian, we can use these conditions to identify critical points that could be local or global optima

• These points are called **KKT points**

• After computing the KKT points, pick the maximizer among them
  • Caution: Remember to verify if the picked point is a maximizer/minimizer
  • How to verify?
    1. Graphical method
    2. If the objection function $f$ is concave and all constraint functions $g_i$ are convex, then the KKT point that you found is a maximizer
Constrained Optimization with Inequality Constraints (KKT conditions)

- Consider the following constrained NLP:

  \[
  \begin{align*}
  &\text{max } x_1 + x_2 \\
  &x_1^2 + x_2^2 \leq 1 \\
  &x_1 \leq \frac{1}{2}
  \end{align*}
  \]

  \[
  \begin{align*}
  f(x) &= x_1 + x_2 \\
  g_1(x) &= x_1^2 + x_2^2, \quad b_1 = 1 \\
  g_2(x) &= x_1, \quad b_2 = \frac{1}{2}
  \end{align*}
  \]

- Algebraic verification:
  - Check if \( f \) is concave and \( g_i \) is convex for every \( i = 1, \ldots, m \)

  \[
  \begin{align*}
  \nabla^2 f &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \Rightarrow \text{all eigenvalues are non-positive, so } f \text{ is concave} \\
  \nabla^2 g_1 &= \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \Rightarrow \text{all eigenvalues are non-negative, so } g_1 \text{ is convex} \\
  \nabla^2 g_2 &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \Rightarrow \text{all eigenvalues are non-negative, so } g_2 \text{ is convex}
  \end{align*}
  \]

  \( \Rightarrow \text{the unique KKT point is a maximizer} \)
Constrained Optimization with Inequality Constraints (KKT conditions)

- Feasible region:
Constrained Optimization with Inequality Constraints (KKT conditions)

- Iso-line:

\[ x_1 + x_2 = 0 \]

Example: graphical verification

\[ \begin{align*}
    \max & \quad x_1 + x_2 \\
    \text{s.t.} & \quad x_1^2 + x_2^2 \leq 1 \\
    & \quad x_1 \leq \frac{1}{2}
\end{align*} \]

Only one KKT point
It turns out to be a global maximum
(verified that it is a maximizer by the graphical method)
**KKT conditions: the significance of $u_i$**

- Recall that one of the KKT conditions is
  \[ u_i(g_i(x) - b_i) = 0 \text{ for } i = 1,2,...,m \]

- This condition means that either:
  - The constraint is satisfied as equality, i.e., $g_i(x) = b_i$, and $u_i$ may take a non-zero value
  - The constraint is a strict inequality, i.e., $g_i(x) < b_i$, and we must have $u_i = 0$
  - That is, this condition is complementary slackness condition!

- So, the value $u_i$ is the **shadow price** of constraint $i$
Constrained Nonlinear Optimization

- We investigated two kinds of Constrained NLP:
  1. All constraints are equality constraints
     Solution method: Lagrange multipliers
  2. All constraints are inequality constraints
     Solution method: KKT conditions
     - Both these methods can be combined to consider both kinds of constraints together

- Finding a KKT point, or a stationary point for the Lagrangian function, does not guarantee that it is a local or global optimum without further analysis

**Theorem.** If $f(x)$ is concave and each $g_i(x)$ is convex, then every KKT point (for a maximization problem) will be a global maximum

- In practice, it is very difficult to apply these methods to large problems, and consequently, nonlinear optimization is a vibrant area of research stemming from Machine Learning applications
NLP: ALGORITHMS

... where we see algorithms to solve single variable unconstrained nonlinear optimization problems
MOTIVATIONS

... why do we even need algorithms?
Why can’t we use the theory?
Motivations for algorithm

• Suppose we want to solve $\max f(x)$
  where $f(x)$ is a differentiable concave function
• If we can find a point $x^*$ such that $f'(x^*) = 0$, then it must be a global maximum

• We may not have an easy way to find such a stationary point
  • If $f(x)$ is a complicated function, there may not be an easy closed form expression for $x^*$
  • E.g., consider $f(x) = xe^{1-x} - x^6$
  • Is this function concave?
A complicated function

\[ f(x) = xe^{1-x} - x^6 \]

Is \( f(x) \) concave?

\[ f'(x) = e^{1-x}(1-x) - 6x^5 \]
\[ f''(x) = e^{1-x}(x-2) - 30x^4 \]
Motivations for algorithm

• Suppose we want to solve \( \max f(x) \)
  where \( f(x) \) is a differentiable concave function
• If we can find a point \( x^* \) such that \( f'(x^*) = 0 \), then it must be a global maximum

• We may not have an easy way to find such a stationary point
  • If \( f(x) \) is a complicated function, there may not be an easy closed form expression for \( x^* \)
  • E.g., consider \( f(x) = xe^{1-x} - x^6 \)
    • This function is concave
      \[
      f'(x) = e^{1-x}(1-x) - 6x^5
      \]

Question: How do we find a root of \( f'(x) = 0 \)?
We will find a value \( x \) that is numerically close to the value of the root