The continuous optimization problem

- In its most general form, a continuous optimization problem may be written

\[
\min_{x \in \mathbb{R}^n} f(x) \quad \text{subject to} \quad c_i(x) = 0, \quad i \in E \\
c_i(x) \geq 0, \quad i \in I
\]

- The function \( f(x) : \mathbb{R}^n \rightarrow \mathbb{R} \) is called the objective function and is assumed to be twice continuously differentiable.
- The vector \( x \) contains the variables to be estimated.
- The functions \( c_i(x) : \mathbb{R}^n \rightarrow \mathbb{R} \) define constraints on the variables.
- The sets \( E \) and \( I \) are index sets for the equality and inequality constraints, respectively.
- A maximization problem is rewritten as

\[
\max_x f(x) \equiv -\min_x -f(x).
\]

Example

- Consider the problem

\[
\min (x_1 - 2)^2 + (x_2 - 1)^2 \\
\text{subject to } x_2 \geq x_1^2, \\
x_1 < 2 - x_2.
\]

- We may rewrite this problem into general form as

\[
\begin{align*}
f(x) &= (x_1 - 2)^2 + (x_2 - 1)^2, \\
x &= \begin{bmatrix} x_1 \\
x_2 \end{bmatrix}, \\
c(x) &= \begin{bmatrix} c_1(x) \\
c_2(x) \end{bmatrix} = \begin{bmatrix} -x_1^2 + x_2 \\
-x_1 - x_2 + 2 \end{bmatrix}, \\
I &= \{1, 2\}, \; E = \{\}
\end{align*}
\]

The parameter space

- The vector \( x \) will be interpreted as a point in \( \mathbb{R}^n \), the parameter space.
- Points that satisfies all constraints are called feasible and belong to the feasible set \( \Omega \) which is a subset of \( \mathbb{R}^n \).
- At a feasible point \( x \), an inequality constraint \( c_i(x) \geq 0 \) is said to be binding or active if \( c_i(x) = 0 \).
- If \( c_i(x) > 0 \), the constraint is nonbinding or inactive.
- Equality constraints are always active.
The parameter space

- The point \( x \) is said to be on the boundary of the constraint if \( c_i(x) = 0 \) and in the interior of the constraint if \( c_i(x) > 0 \).
- Equality constraints have no interior points.
- The set of active constraints at a given point is called the active set (of constraints).
- A feasible point with at least one active constraint belongs to the boundary of the feasible set.
- Feasible points with no active constraints are interior points to the feasible set.

Optimality conditions for constrained problems

- A minimizer \( x^* \) to a minimization problem
  \[
  \min_x f(x) \\
  \text{s.t. } c_i(x) = 0, \quad i \in \mathcal{E} \\
  c_i(x) \geq 0, \quad i \in \mathcal{I}
  \]
  must satisfy
  \[
  \rho^T \nabla f(x^*) \geq 0
  \]
  for all feasible directions \( \rho \).

Overview

- Consider a problem with linear equality constraints, i.e.
  \[
  \min_x f(x) \\
  \text{s.t. } Ax = b,
  \]
  where \( A \) is assumed to have full rank.
- The constrained problem may be rewritten to the unconstrained problem
  \[
  \min_{\nu \in \mathbb{R}^r} \phi(\nu) = f(\bar{x} + Z \nu),
  \]
  where \( \bar{x} \) is a feasible point and \( Z \in \mathbb{R}^{n \times r} \) is a basis for \( \mathcal{N}(A) \).
- The function \( \phi(\nu) \) is called the reduced function.
Necessary conditions for the reduced problem

▶ The necessary conditions for the reduced problem are

\[
\nabla \phi(v) = Z^T \nabla f(\bar{x} + Zv) = Z^T \nabla f(x) = 0,
\]

\[
\nabla^2 \phi(v) = Z^T \nabla^2 f(\bar{x} + Zv)Z = Z^T \nabla^2 f(x)Z \quad \text{pos. semidef.,}
\]

where \( x = \bar{x} + Zv \).

▶ The expression

\[
Z^T \nabla f(x)
\]

is called the reduced gradient and

\[
Z^T \nabla^2 f(x)Z
\]

the reduced Hessian.

▶ If the null space matrix \( Z \) is an orthogonal projection matrix, they are called projected gradient and Hessian, respectively.

Sufficient conditions for linear equality constraints

▶ If \( x^* \) satisfies

\[
A x^* = b,
\]

\[
Z^T \nabla f(x^*) = 0, \quad \text{and}
\]

\[
Z^T \nabla^2 f(x^*)Z \quad \text{is positive definite},
\]

where \( Z \) is a basis matrix for \( \mathcal{N}(A) \), then \( x^* \) is a strict local minimizer of \( f \) over \( \{ x : Ax = b \} \).

Example

▶ Consider the problem

\[
\min_{x} f(x) = x_1^2 - 2x_1 + x_2^2 - x_3^2 + 4x_3,
\]

s.t. \( x_1 - x_2 + 2x_3 = 2 \),

with gradient and Hessian functions

\[
\nabla f(x) = \begin{bmatrix} 2x_1 - 2 \\ 2x_2 \\ -2x_3 + 4 \end{bmatrix} \quad \text{and} \quad \nabla^2 f(x) = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{bmatrix}.
\]

▶ As null space matrix of the constraint matrix

\[
A = [1, -1, 2]
\]

we may choose

\[
Z = \begin{bmatrix} 1 & -2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}.
\]
Example: The reduced gradient

- In the feasible point
  \[ x = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}, \]
  the reduced gradient is
  \[ Z^T \nabla f(x) = \begin{bmatrix} 1 & -2 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}^T \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \]
  and \( x \) is not a local minimizer.

- In the feasible point
  \[ x^* = \begin{bmatrix} 2.5 \\ -1.5 \\ -1 \end{bmatrix}, \nabla f(x^*) = \begin{bmatrix} 3 \\ -3 \\ -6 \end{bmatrix}, \text{ and } Z^T \nabla f(x) = 0. \]
  Hence, \( x^* \) is potentially a local minimizer.

The Lagrange multipliers

- Let \( x^* \) be a minimizer and \( Z \in \mathbb{R}^{n \times r} \) a null space matrix for \( A \).
- The gradient \( \nabla f(x^*) \) may be expressed as the sum of its null space and range space components, i.e.
  \[ \nabla f(x^*) = Zv^* + A^T \lambda^* , \]
  where \( v^* \in \mathbb{R}^r \) and \( \lambda^* \in \mathbb{R}^m \).
- Together with the first order conditions we get
  \[ Z^T \nabla f(x^*) = Z^T Zv^* + Z^T A^T \lambda^* = \]
  \[ = Z^T Zv^* + (AZ)^T \lambda^* = 0 \]
  \[ = Z^T Zv^* = 0 \]
  \[ \Downarrow \]
  \[ Zv^* = 0. \]

Example: The reduced Hessian

- The reduced Hessian in \( x^* \) is
  \[ Z^T \nabla^2 f(x^*) Z = \begin{bmatrix} 1 & -2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}^T \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 4 & \lambda^* \end{bmatrix} \]
  and is positive definite.
- Thus, the second order sufficient conditions are satisfied and \( x^* \) is a strict local minimizer of \( f \).
- Notice that \( \nabla^2 f(x) \) itself is not positive definite.

The Lagrange multipliers

- Thus, a minimizer satisfies
  \[ \nabla f(x^*) = A^T \lambda^* . \]
- In other words: In a local minimum, the gradient of the objective function is a linear combination of the gradients of the constraints.
- The coefficients in the vector \( \lambda^* \) are called Lagrange multipliers.
- The constraint and first order condition may be formulated in one system equation of \( n + m \) equations and \( n + m \) unknowns in \( x \) and \( \lambda \):
  \[ \begin{bmatrix} 0 & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} \nabla f(x) \\ b \end{bmatrix} . \]
The Lagrange multipliers and sensitivity

- The Lagrange multipliers may be used to estimate the sensitivity of the min value $f(x^*)$ with respect to the constraints.
- Assume we have found a solution $x^*$ to
  \[
  \min_x f(x) \quad \text{s.t.} \quad Ax = b.
  \]
- Consider the perturbed constraints $Ax = b + \delta$.

The Lagrangian function

- Define the Lagrangian function of $x$ and $\lambda$ as
  \[
  \mathcal{L}(x, \lambda) = f(x) - \sum_{i=1}^m \lambda_i (a_i^T x - b_i) = f(x) - \lambda^T (Ax - b).
  \]
- The gradients of the Lagrangian are
  \[
  \nabla_x \mathcal{L}(x, \lambda) = \nabla f(x) - A^T \lambda
  \]
  and
  \[
  \nabla_{\lambda} \mathcal{L}(x, \lambda) = b - Ax.
  \]
- The first order condition on the Lagrangian
  \[
  \nabla \mathcal{L}(x^*, \lambda^*) = 0
  \]
correspond to the first order condition on the constrained problem.
- A local minimizer to the constrained problem is a stationary point to the Lagrangian.

The Lagrange multipliers and sensitivity

- The perturbation $\delta$ is small enough, the solution $\bar{x}$ to the perturbed problem will be close to $x^*$ and
  \[
  f(\bar{x}) \approx f(x^*) + (\bar{x} - x^*)^T \nabla f(x^*) = f(x^*) + (\bar{x} - x^*)^T A^T \lambda^*
  \]
  \[
  = f(x^*) + (Ax - Ax^*)^T \lambda^* = f(x^*) + (b - \delta - b)^T \lambda^*
  \]
  \[
  = f(x^*) + \delta^T \lambda^*
  \]
  \[
  = f(x^*) + \sum_{i=1}^m \delta_i \lambda_{\delta i}.
  \]
- Thus, if element $i$ of the right hand side of the constraint is modified by $\delta_i$, the optimal objective value will change with about $\delta_i \lambda_{\delta i}$.
- For this reason, the Lagrange multipliers are sometimes called shadow prices or dual variables.

Example

\[
\begin{align*}
A &= \begin{bmatrix} 0 & -1 \end{bmatrix}, \quad Z = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \\
x^* &= \begin{bmatrix} 0 \\ 0.6 \end{bmatrix}, \quad \nabla f(x^*) = \begin{bmatrix} 0 \\ 1.35 \end{bmatrix}, \\
\nabla^2 f(x^*) &= \begin{bmatrix} 2 & 0 \\ 0 & -5.9 \end{bmatrix}, \\
Z^T \nabla^2 f(x^*) Z &= [2].
\end{align*}
\]
Optimality conditions for linear inequality constrained problems

- Consider a problem with linear inequality constraints, i.e.

\[
\begin{align*}
\min_{x} \quad & f(x) \\
\text{s.t.} \quad & Ax \geq b,
\end{align*}
\]

where \( A \) is assumed to have full rank.

- The active constraints in a point \( x^* \) will determine if \( x^* \) is a minimizer.

- Our problem may thus be rewritten as

\[
\begin{align*}
\min_{x} \quad & f(x) \\
\text{s.t.} \quad & \hat{A}x = \hat{b},
\end{align*}
\]

where \( \hat{A} \) and \( \hat{b} \) contains the active constraints.

Complementary slackness

- If we define the Lagrange multiplier of an inactive constraint to be zero, we may describe the inequality conditions as

\[
\lambda_i (a_i^T x^* - b_i) = 0, \quad i = 1, \ldots, m.
\]

- This condition is called complementary slackness and means that
  - either the constraint is active (\( a_i^T x^* - b_i = 0 \))
  - or the Lagrange multiplier is zero (\( \lambda_i = 0 \)).

- At least one of the two must be true.

- The case when both cannot be true at the same time is called strict complementarity.

- Without strict complementarity, a Lagrange multiplier may be zero even if the constraint is active.

- In such a case, that constraint is called degenerate.

Example

For the problem

\[
\begin{align*}
\min_{x} \quad & f(x) = x^2 + \sin^2 2y \\
\text{s.t.} \quad & \begin{bmatrix} 0 & -1 \\ 0 & 1 \\ -2 & 1 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \geq \begin{bmatrix} -1 \\ -1 \end{bmatrix}
\end{align*}
\]

there are four corners, two of which are degenerate.

<table>
<thead>
<tr>
<th>Point</th>
<th>Active constraints</th>
<th>( (x, y) )</th>
<th>( \nabla f )</th>
<th>( \lambda )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( q_1 )</td>
<td>( 1, 3 )</td>
<td>( 0 )</td>
<td>( 0 )</td>
<td>( 1.5, 0 )</td>
</tr>
<tr>
<td>( q_2 )</td>
<td>( 1, 4 )</td>
<td>( -0.5 )</td>
<td>( -1 )</td>
<td>( 1.5, -0.5 )</td>
</tr>
<tr>
<td>( q_3 )</td>
<td>( 2, 3 )</td>
<td>( 0.5 )</td>
<td>( 1 )</td>
<td>( -0.5, -0.5 )</td>
</tr>
<tr>
<td>( q_4 )</td>
<td>( 2, 4 )</td>
<td>( -0.5 )</td>
<td>( 0 )</td>
<td>( 0.5, -0.5 )</td>
</tr>
</tbody>
</table>
Necessary condition, linear inequalities

- In summary, the following conditions have to be satisfied for a point $x^*$ to be a minimizer of a function $f$ on the set $\{x : Ax \geq b\}$:
  
  $\begin{align*}
  Ax^* &\geq b, \\
  \nabla f(x^*) &= A^T \lambda^* \iff Z^T \nabla f(x^*) = 0, \\
  \lambda^* &\geq 0, \\
  \lambda^*^T (Ax^* - b) &= 0, \\
  Z^T \nabla^2 f(x^*) Z &\text{ positive semidefinite,}
  \end{align*}$

  for some vector $\lambda^*$ of Lagrange multipliers and where $Z$ is a null space matrix for the matrix $\hat{A}$ of the active constraints in $x^*$.

Sufficient conditions, linear inequalities I

- With strict complementarity we may extend to sufficient conditions in a straightforward manner:
  
  - Assume $x^*$ satisfies
    
    $\begin{align*}
    Ax^* &\geq b, \\
    \nabla f(x^*) &= A^T \lambda^*, \\
    \lambda^* &\geq 0, \\
    \lambda^*^T (Ax^* - b) &= 0, \\
    Z^T \nabla^2 f(x^*) Z &\text{ positive definite,}
    \end{align*}$

  - Then $x^*$ is a strict local minimizer of $f$ on the set $\{x : Ax \geq b\}$.

Why strict complementarity is needed

- The point $x^*$ is also a strict local minimizer on the set $\{x : \hat{A}x = \hat{b}\}$, i.e. $f$ increases in all directions $p$ such that $\hat{A}p = 0$:
  
  - Study a direction $p$ such that $\hat{A}p \geq 0$, where some component of $p$ is strictly positive, i.e. $p$ points into the feasible set.
  
  - Since $\nabla f(x^*) = A^T \lambda^* = \hat{A}^T \hat{\lambda}_*$, then $p^T \nabla f(x^*) = p^T \hat{A}^T \hat{\lambda}_* = (\hat{A}p)^T \hat{\lambda}_*$.

Why strict complementarity is needed

Cont’d

- With strict complementarity, we know that $$(\hat{A}p)^T \hat{\lambda}_* > 0,$$ i.e. $p$ is an ascent direction and $x^*$ must be a strict minimizer.
  
  - Without strict complementarity, $$(\hat{A}p)^T \hat{\lambda}_* = 0$$ may be true for some $p$.
  
  - This, we cannot tell anything about $x^*$ with only first order information.
  
  - However, if we drop the degenerate constraints, we may formulate sufficient conditions on the remaining constraints.
Sufficient conditions, linear inequalities II

Let \( \hat{A}_+ \) contain the rows of \( \hat{A} \) corresponding to the non-degenerate constraints in \( x^* \).

Let \( Z_+ \) be a null space matrix to \( \hat{A}_+ \).

Assume \( x^* \) satisfies

\[
\begin{align*}
Ax^* & \geq b, \\
\nabla f(x^*) &= A^T \lambda^*, \\
\lambda^* & \geq 0, \\
\lambda^*^T (Ax^* - b) &= 0, \\
Z_+^T \nabla^2 f(x^*) Z_+ & \text{ positive definite.}
\end{align*}
\]

Then \( x^* \) is a strict local minimizer to the inequality constrained problem.

Optimality conditions for non-linear constraints

- Non-linear optimization problems with non-linear constraints are formulated as

\[
\begin{align*}
\min_{x} & \quad f(x) \\
\text{s.t.} & \quad c_i(x) = 0, i = 1, \ldots, m
\end{align*}
\]

for equality constraints, and

\[
\begin{align*}
\min_{x} & \quad f(x) \\
\text{s.t.} & \quad c_i(x) \geq 0, i = 1, \ldots, m
\end{align*}
\]

for inequality constraints.

- We will assume that the solution point \( x^* \) is regular, i.e. that the gradients of the active constraints in \( x^* \) \{\( \nabla c_i(x^*) : c_i(x^*) = 0 \}\) are linearly independent.

Example

For the problem

\[
\begin{align*}
\min_{x} & \quad f(x) = x^2 + \sin^2 2y \\
\text{s.t.} & \quad \begin{bmatrix} 0 & -1 \\ 0 & 1 \\ -2 & -1 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \geq \begin{bmatrix} -1 \\ 0 \\ -1 \\ -1 \end{bmatrix}
\end{align*}
\]

we have

\[
\begin{align*}
\lambda_1 &= 1.5, 0 \\
\lambda_2 &= 1.5, -0.5 \\
\lambda_3 &= -0.5, -0.5 \\
\lambda_4 &= 0, -0.5
\end{align*}
\]

Optimality conditions for non-linear constraints

Cont’d

- The optimality conditions are expressed in terms of the Lagrangian function

\[
\mathcal{L}(x, \lambda) = f(x) - \sum_{i=1}^{m} \lambda_i c_i(x) = f(x) - \lambda^T c(x),
\]

where \( \lambda \) is a vector of Lagrange multipliers and \( c \) is a vector of constraint functions \( \{c_i\} \).
Necessary conditions for equality constraints

- Let \( x^* \) be a local minimizer for \( f \) under the constraints \( c(x) = 0 \) and \( Z(x^*) \) be a null space matrix for the Jacobian \( \nabla c(x^*)^T \) of the constraints.
- If \( x^* \) is a regular point, then there exists a Lagrangian vector \( \lambda^* \) such that

\[
\begin{align*}
\nabla_x L(x^*, \lambda^*) &= 0 \iff Z(x^*)^T \nabla f(x^*) = 0, \\
Z(x^*)^T \nabla^2_{xx} L(x^*, \lambda^*) Z(x^*) &\text{ positive semi-definite.}
\end{align*}
\]

Sufficient conditions for equality constraints

- Let \( x^* \) be a point such that \( c(x^*) = 0 \) and \( Z(x^*) \) is a basis for the null space of \( \nabla c(x^*)^T \).
- Assume there exists a vector \( \lambda^* \) such that

\[
\begin{align*}
\nabla_x L(x^*, \lambda^*) &= 0, \\
Z(x^*)^T \nabla^2_{xx} L(x^*, \lambda^*) Z(x^*) &\text{ positive definite.}
\end{align*}
\]
- Then \( x^* \) is a strict local minimizer to \( f \) on the constraint set \( \{ x : c(x) = 0 \} \).

Example

- For linear constraints \( c(x) = Ax - b \), the Jacobian is \( \nabla c(x)^T = A \) and the first order conditions

\[
Z(x^*)^T \nabla f(x^*) = 0 \iff \nabla_x L(x^*, \lambda^*) = \nabla f(x^*) - \nabla c(x^*) \lambda = 0
\]

becomes

\[
Z^T \nabla f(x^*) = 0 \iff \nabla f(x^*) = A^T \lambda^*.
\]

- The second order necessary (sufficient) conditions becomes that

\[
\nabla^2_{xx} L(x^*, \lambda^*) = \nabla^2 f(x^*)
\]

should be positive semi-definite (definite).
Necessary conditions for inequality constraints

- Let \( x^* \) be a local minimizer for \( f \) under the constraints \( c(x) \geq 0 \) and \( Z(x^*) \) be a null space matrix for the Jacobian of the active constraints in \( x^* \).
- If \( x^* \) is a regular point, then there exists a Lagrangian vector \( \lambda^* \) such that
  \[
  \nabla_x L(x^*, \lambda^*) = 0 \iff Z(x^*)^T \nabla f(x^*) = 0, \\
  \lambda^* \geq 0, \\
  \lambda^*^T c(x^*) = 0, \\
  Z(x^*)^T \nabla^2_{xx} L(x^*, \lambda^*) Z(x^*) \text{ positive semi-definite.}
  \]
- The condition \( \lambda^*^T c(x^*) = 0 \) is the non-linear version of the complementary slackness condition.

Sufficient conditions for inequality constraints

- Let \( x^* \) be a points such that \( c(x^*) \geq 0 \).
- Assume there exists a vector \( \lambda^* \) such that
  \[
  \nabla_x L(x^*, \lambda^*) = 0, \\
  \lambda^* \geq 0, \\
  \lambda^*^T c(x^*) = 0, \\
  Z_+(x^*)^T \nabla^2_{xx} L(x^*, \lambda^*) Z_+(x^*) \text{ is positive definite,}
  \]
  where \( Z_+(x^*) \) is a basis for the null space of the Jacobian of the non-degenerate constraints in \( x^* \).
- Then \( x^* \) is a strict local minimizer to \( f \) on the constraint set \( \{ x : c(x) \geq 0 \} \).
- The necessary and sufficient conditions for the non-linear inequality constraints are often called the KKT conditions (Karush-Kuhn-Tucker conditions).

Duality

- The concept of duality is that for each minimization problem, there is a corresponding maximization problem that under some circumstances both problems have the same optimum.
- Define
  \[
  F^*(x) = \max_{y \in Y} F(x, y) \\
  F_*(y) = \min_{x \in X} F(x, y).
  \]

Duality Cont’d

- The problem \( \min_{x \in X} F^*(x) = \min_{x \in X} \max_{y \in Y} F(x, y) \) is called a min-max problem and the problem \( \max_{y \in Y} F_*(y) = \max_{y \in Y} \min_{x \in X} F(x, y) \) is called a max-min problem.
- These problems are each others duals.
- The min-max problem is called the primal problem and \( F^*(x) \) is called the primal function.
- The max-min problem is called the dual problem and \( F_*(y) \) is called the dual function.
Weak duality

- Each \( x \in X \) and \( y \in Y \) satisfies
  \[
  F_*(y) = \min_{x \in X} F(x, y) \leq F(x, y) \leq \max_{y \in Y} F(x, y) = F^*(x)
  \]
or
  \[
  F_*(y) \leq F^*(x).
  \]

- This is called weak duality.
- A consequence of weak duality is that the primal problem is bounded from below by \( F_*(y) \).

Strong duality

- A point \((x^*, y_*)\) satisfies the saddle-point condition for \( F \) if
  \[
  F(x^*, y) \leq F(x^*, y_*) \leq F(x, y_*)
  \]
  for all \( x \in X \) and \( y \in Y \).
- Assume there exists a point \((x^*, y_*)\) that satisfies the saddle-point condition.
- Then the solution value of the primal and the dual problem is the same, i.e.
  \[
  \min_{x \in X} \max_{y \in Y} F(x, y) = \max_{y \in Y} \min_{x \in X} F(x, y).
  \]

- This is called strong duality.

Duality and the Lagrange multipliers

- Consider the non-linear problem
  \[
  \begin{align*}
  \min_x & \quad f(x) \\
  \text{s.t.} & \quad c_i(x) \geq 0, \quad i = 1, \ldots, m
  \end{align*}
  \]
  and its corresponding Lagrangian function
  \[
  L(x, \lambda) = f(x) - \lambda^T c(x),
  \]
  where \( x \in \mathbb{R}^n \), \( \lambda \in \mathbb{R}^m \), \( \lambda \geq 0 \).
- Define the primal function
  \[
  L^*(x) = \max_{\lambda \geq 0} L(x, \lambda).
  \]

Duality and the Lagrange multipliers

Cont’d

- Study \( L^*(x) \) for a fixed \( x \):
  \[
  L^*(x) = \max_{\lambda \geq 0} \left( f(x) - \lambda^T c(x) \right).
  \]
- For a feasible point, \( c(x) \geq 0 \) and \( L^*(x) = f(x) \).
- For an infeasible point, some constraint \( c_i(x) \) will be negative and \( L^*(x) \) will be without bound.
- If we formulate the primal problem as
  \[
  \min_x L^*(x),
  \]
  then it will be the same as our original constrained problem.
Duality and the Lagrange multipliers

- We may use min-max-duality to formulate the dual problem.
- For each $\lambda \geq 0$, define the dual function
  \[ \mathcal{L}^*(\lambda) = \min_x \mathcal{L}(x, \lambda) \]
  and the dual max-min-problem
  \[ \max_{\lambda \geq 0} \mathcal{L}^*(\lambda). \]
- Some methods work on the dual problem instead of the primal.