MATH2070 Optimisation

Nonlinear optimisation with constraints

Semester 2, 2012
Lecturer: I.W. Guo
Lecture slides courtesy of J.R. Wishart
Review

The full nonlinear optimisation problem with equality constraints

Method of Lagrange multipliers

Dealing with Inequality Constraints and the Kuhn-Tucker conditions

Second order conditions with constraints
The full nonlinear optimisation problem with equality constraints

Method of Lagrange multipliers

Dealing with Inequality Constraints and the Kuhn-Tucker conditions

Second order conditions with constraints
Non-linear optimisation

Interested in optimising a function of several variables with constraints.

### Multivariate framework

<table>
<thead>
<tr>
<th>Variables</th>
<th>$\mathbf{x} = (x_1, x_2, \ldots, x_n) \in D \subset \mathbb{R}^n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Objective Function</td>
<td>$Z = f(x_1, x_2, \ldots, x_n)$</td>
</tr>
<tr>
<td>Constraint equations,</td>
<td>$g_1(x_1, x_2, \ldots, x_n) = 0, \quad h_1(x_1, x_2, \ldots, x_n) \leq 0,$</td>
</tr>
<tr>
<td></td>
<td>$g_2(x_1, x_2, \ldots, x_n) = 0, \quad h_2(x_1, x_2, \ldots, x_n) \leq 0,$</td>
</tr>
<tr>
<td></td>
<td>$\vdots$</td>
</tr>
<tr>
<td></td>
<td>$g_m(x_1, x_2, \ldots, x_n) = 0, \quad h_p(x_1, x_2, \ldots, x_n) \leq 0.$</td>
</tr>
</tbody>
</table>
## Vector notation

### Multivariate framework

Minimise \[ Z = f(x) \]

such that \[ g(x) = 0, \quad h(x) \leq 0, \]

where \( g : \mathbb{R}^n \rightarrow \mathbb{R}^m \) and \( h : \mathbb{R}^n \rightarrow \mathbb{R}^p \).

### Objectives?

- Find extrema of above problem.
- Consider equality constraints first.
- Extension: Allow inequality constraints.
Initial Example

Minimise the heat loss through the surface area of an open lid rectangular container.

Formulated as a problem

Minimise,

\[ S = 2xy + 2yz + xz \]

such that,

\[ V = xyz \equiv \text{constant} \quad \text{and} \quad x, y, z > 0. \]

- Parameters : \((x, y, z)\)
- Objective function : \(f = S\)
- Constraints : \(g = xyz - V\)
Graphical interpretation of problem

- Feasible points occur at the tangential intersection of \( g(x, y) \) and \( f(x, y) \).
- Gradient of \( f \) is parallel to gradient of \( g \).
- Recall from linear algebra and vectors

\[
\nabla f + \lambda \nabla g = 0,
\]

for some constant, \( \lambda \).
Lagrangian

**Definition**

Define the *Lagrangian* for our problem,

\[
L(x, y, z, \lambda) = S(x, y, z) + \lambda g(x, y, z)
\]

\[
= 2xy + 2yz + xz + \lambda (xyz - V)
\]

- **Necessary conditions,**

\[
\frac{\partial L}{\partial x} = \frac{\partial L}{\partial y} = \frac{\partial L}{\partial z} = \frac{\partial L}{\partial \lambda} = 0.
\]

- **This will be a system of four equations with four unknowns** \((x, y, z, \lambda)\) that needs to be solved
Solving the system

- Calculate derivatives,

\[ L_x = 2y + z - \lambda yz, \quad L_y = 2x + 2z - \lambda xz, \quad L_z = 2y + x - \lambda xy, \]
\[ L_\lambda = xyz - V, \]

- Ensure constraint,

\[ L_\lambda = 0 \Rightarrow z = \frac{V}{xy} \quad (1) \]

- Eliminate \( \lambda \), solve for \((x, y, z)\)

\[ L_x = 0 \Rightarrow \quad 2y + z = \lambda yz \Rightarrow \quad \lambda = \frac{2}{z} + \frac{1}{y} \quad (2) \]
\[ L_y = 0 \Rightarrow \quad 2x + 2z = \lambda xz \Rightarrow \quad \lambda = \frac{2}{z} + \frac{2}{x} \quad (3) \]
\[ L_z = 0 \Rightarrow \quad 2y + 2x = \lambda xy \Rightarrow \quad \lambda = \frac{2}{x} + \frac{1}{y} \quad (4) \]
Solving system (continued)

- From (2) and (3) it follows $x = 2y$

- From (2) and (4) it follows $x = z$

- Use (1) with $x = z = 2y$ then simplify,

$$\mathbf{x} = (x, y, z) = \left( \sqrt[3]{2V}, \sqrt[3]{\frac{V}{4}}, \sqrt[3]{2V} \right)$$
The full nonlinear optimisation problem with equality constraints

Method of Lagrange multipliers

Dealing with Inequality Constraints and the Kuhn-Tucker conditions

Second order conditions with constraints
Constrained optimisation

Wish to generalised previous argument to multivariate framework,

Minimise \( Z = f(x), \ x \in D \subset \mathbb{R}^n \)

such that \( g(x) = 0, \ h(x) \leq 0, \)

where \( g: \mathbb{R}^n \rightarrow \mathbb{R}^m \) and \( h: \mathbb{R}^n \rightarrow \mathbb{R}^p. \)

- Geometrically, the constraints define a subset \( \Omega \) of \( \mathbb{R}^n \), the feasible set, which is of dimension less than \( n - m \), in which the minimum lies.
Constrained Optimisation

- A value $x \in D$ is called **feasible** if both $g(x) = 0$ and $h(x) \leq 0$.

- An inequality constraint $h_i(x) \leq 0$ is
  - Active if the feasible $x \in D$ satisfies $h_i(x) = 0$
  - Inactive otherwise.
Equality Constraints

A point \( x^* \) which satisfies the constraint \( g(x^*) = 0 \) is a **regular point** if

\[
\{ \nabla g_1(x^*), \nabla g_2(x^*), \ldots, \nabla g_m(x^*) \}
\]

is linearly independent.

\[
\text{Rank } \begin{pmatrix}
\frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\
\vdots & \ddots & \vdots \\
\frac{\partial g_m}{\partial x_1} & \cdots & \frac{\partial g_m}{\partial x_n}
\end{pmatrix} = m.
\]
Tangent Subspace

**Definition**

Define the constrained tangent subspace

\[
\mathcal{T} = \{ d \in \mathbb{R}^n : \nabla g(x^*)d = 0 \},
\]

where \( \nabla g(x^*) \) is an \( m \times n \) matrix defined,

\[
\nabla g(x^*) := \begin{pmatrix}
\frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\
\vdots & \ddots & \vdots \\
\frac{\partial g_m}{\partial x_1} & \cdots & \frac{\partial g_m}{\partial x_n}
\end{pmatrix} = \begin{pmatrix}
\nabla g_1(x^*)^T \\
\vdots \\
\nabla g_m(x^*)^T
\end{pmatrix}
\]

and \( \nabla g_i^T := \left( \frac{\partial g_i}{\partial x_1}, \ldots, \frac{\partial g_i}{\partial x_n} \right) \)
Tangent Subspace (Example)
Generalised Lagrangian

**Definition**

Define the Lagrangian for the objective function \( f(x) \) with constraint equations \( g(x) = 0 \),

\[
L(x, \lambda) := f + \sum_{i=1}^{m} \lambda_i g_i = f + \lambda^T g.
\]

▶ Introduce a lagrange multiplier for each equality constraint.
First order necessary conditions

**Theorem**

Assume $\mathbf{x}^*$ is a regular point of $g(\mathbf{x}) = 0$ and let $\mathbf{x}^*$ be a local extremum of $f$ subject to $g(\mathbf{x}) = 0$, with $f, g \in C^1$. Then there exists $\lambda \in \mathbb{R}^m$ such that,

$$\nabla L(\mathbf{x}^*, \lambda) = \nabla f(\mathbf{x}^*) + \lambda^T \nabla g(\mathbf{x}^*) = 0.$$

In sigma notation:

$$\frac{\partial L(\mathbf{x}^*)}{\partial x_j} = \frac{\partial f(\mathbf{x}^*)}{\partial x_j} + \sum_{i=1}^{m} \lambda_i \frac{\partial g_i(\mathbf{x}^*)}{\partial x_j} = 0, \quad j = 1, \ldots, n;$$

and

$$\frac{\partial L(\mathbf{x}^*)}{\partial \lambda_i} = g_i(\mathbf{x}^*) = 0, \quad i = 1, \ldots, m.$$
Another example

Example

Find the extremum for \( f(x, y) = \frac{1}{2}(x - 1)^2 + \frac{1}{2}(y - 2)^2 + 1 \) subject to the single constraint \( x + y = 1 \).

Solution: Construct the Lagrangian:

\[
L(x, y, \lambda) = \frac{1}{2}(x - 1)^2 + \frac{1}{2}(y - 2)^2 + 1 + \lambda(x + y - 1).
\]

Then the necessary conditions are:

\[
\frac{\partial L}{\partial x} = x - 1 + \lambda = 0 \\
\frac{\partial L}{\partial y} = y - 2 + \lambda = 0 \\
\frac{\partial L}{\partial \lambda} = 0 \Rightarrow x + y = 1.
\]
Solve system

Can solve this by converting problem into matrix form.

i.e. solve for $a$ in $Ma = b$

\[
\begin{pmatrix}
1 & 0 & 1 \\
0 & 1 & 1 \\
1 & 1 & 0
\end{pmatrix}
\begin{pmatrix}
x \\
y \\
\lambda
\end{pmatrix}
= 
\begin{pmatrix}
1 \\
2 \\
1
\end{pmatrix}
\]

Invert matrix $M$ and solve gives,

\[
\begin{pmatrix}
x \\
y \\
\lambda
\end{pmatrix}
= 
\begin{pmatrix}
1 & 0 & 1 \\
0 & 1 & 1 \\
1 & 1 & 0
\end{pmatrix}^{-1}
\begin{pmatrix}
1 \\
2 \\
1
\end{pmatrix}
= 
\begin{pmatrix}
0 \\
1 \\
1
\end{pmatrix}
\]
Comparison with unconstrained problem

Example

Minimize $f(x, y) = \frac{1}{2}(x - 1)^2 + \frac{1}{2}(y - 2)^2 + 1$ subject to the single constraint $x + y = 1$.

- In unconstrained problem, minimum attained at $x = (1, 2)$.
  
  Minimum at $f^* = f(1, 2) = 1$

- Constraint modifies solution.

  Constrained minimum at $f^* = f(0, 1) = 2$.

- Note: Constrained minimum is bounded from below by the unconstrained minimum.
Use of convex functions: Example

Example

Minimise: $f(x, y) = x^2 + y^2 + xy - 5y$ such that: $x + 2y = 5$.

The constraint is linear (therefore convex) and Hessian matrix of $f$ is positive definite so $f$ is convex as well.

The Lagrangian is

$$L(x, y, \lambda) = x^2 + y^2 + xy - 5y + \lambda(x + 2y - 5).$$
Finding first order necessary conditions,
\[
\frac{\partial L}{\partial x} = 2x + y + \lambda = 0 \tag{5}
\]
\[
\frac{\partial L}{\partial y} = 2y + x - 5 + 2\lambda = 0 \tag{6}
\]
\[
\frac{\partial L}{\partial \lambda} = x + 2y - 5 = 0 \tag{7}
\]
In matrix form the system is,
\[
\begin{pmatrix}
2 & 1 & 1 \\
1 & 2 & 2 \\
1 & 2 & 0
\end{pmatrix}
\begin{pmatrix}
x \\
y \\
\lambda
\end{pmatrix}
= 
\begin{pmatrix}
0 \\
5 \\
5
\end{pmatrix}
\]
Inverting and solving yields,
\[
\begin{pmatrix}
x \\
y \\
\lambda
\end{pmatrix}
= 
\frac{1}{3}
\begin{pmatrix}
-5 \\
10 \\
0
\end{pmatrix}
\]
The full nonlinear optimisation problem with equality constraints

Method of Lagrange multipliers

Dealing with Inequality Constraints and the Kuhn-Tucker conditions

Second order conditions with constraints
Inequality constraints

We now consider constrained optimisation problems with \( p \) inequality constraints,

Minimise: \( f(x_1, x_2, \ldots, x_n) \)
subject to: \( h_1(x_1, x_2, \ldots, x_n) \leq 0 \)
\( h_2(x_1, x_2, \ldots, x_n) \leq 0 \)
\[ \vdots \]
\( h_p(x_1, x_2, \ldots, x_n) \leq 0 \).

In vector form the problem is

Minimise: \( f(x) \)
subject to: \( h(x) \leq 0. \)
The Kuhn-Tucker Conditions

- To deal with the inequality constraints, introduce the so called Kuhn-Tucker conditions.

\[
\frac{\partial f}{\partial x_j} + \sum_{i=1}^{p} \lambda_i \frac{\partial h_i}{\partial x_j} = 0, \quad j = 1, \ldots, n;
\]
\[
\lambda_i h_i = 0, \quad i = 1, \ldots, m;
\]
\[
\lambda_i \geq 0, \quad i = 1, \ldots, m.
\]

- Extra conditions for \( \lambda \) apply in this sense,

\[ h_i(x) = 0, \quad \text{if } \lambda_i > 0, \ i = 1, \ldots, p \]

or

\[ h_i(x) < 0, \quad \text{if } \lambda_i = 0, \ i = 1, \ldots, p. \]
Graphical Interpretation

- If $\lambda > 0$ then find solution on line $h(x, y) = 0$
- If $\lambda = 0$ then find solution in region $h(x, y) < 0$
Graphical Interpretation

Consider $h(x, y) = 0$, then $\lambda > 0$. Optimum at $x^* = (x^*, y^*)$
Graphical Interpretation

Consider $h(x, y) < 0$, then $\lambda = 0$. Optimum at $x^* = (x^*, y^*)$
Kuhn Tucker

- Kuhn-Tucker conditions are always necessary.
- Kuhn-Tucker conditions are sufficient if the objective function and the constraint functions are convex.
- Need a procedure to deal with these conditions.
Introduce slack variables

Slack Variables

To deal with the inequality constraint introduce a variable, $s^2_i$, for each constraint.

New constrained optimisation problem with $m$ equality constraints,

Minimise: $f(x_1, x_2, \ldots, x_n)$

subject to: $h_1(x_1, x_2, \ldots, x_n) + s^2_1 = 0$
$h_2(x_1, x_2, \ldots, x_n) + s^2_2 = 0$
$\vdots$
$h_p(x_1, x_2, \ldots, x_n) + s^2_p = 0$. 

New Lagrangian

The Lagrangian for constrained problem is,

\[ L(x, \lambda, s) = f(x) + \sum_{i=1}^{p} \lambda_i \left( h_i(x) + s_i^2 \right) \]

\[ = f(x) + \lambda^T h(x) + \lambda^T s, \]

where \( s \) is a vector in \( \mathbb{R}^p \) with \( i \)-th component \( s_i^2 \).

- \( L \) depends on the \( n + 2m \) independent variables \( x_1, x_2, \ldots, x_n, \lambda_1, \lambda_2, \ldots, \lambda_p \) and \( s_1, s_2, \ldots, s_p \).
- To find the constrained minimum we must differentiate \( L \) with respect to each of these variables and set the result to zero in each case.
Solving the system

Need to solve \( n + 2m \) variables given by the \( n + 2m \) equations generated by,

\[
\frac{\partial L}{\partial x_1} = \frac{\partial L}{\partial x_2} = \cdots = \frac{\partial L}{\partial x_n} = 0
\]

\[
\frac{\partial L}{\partial \lambda_1} = \frac{\partial L}{\partial \lambda_2} = \cdots = \frac{\partial L}{\partial \lambda_p} = \frac{\partial L}{\partial s_1} = \frac{\partial L}{\partial s_2} = \cdots = \frac{\partial L}{\partial s_p} = 0.
\]

- Having \( \lambda_i \geq 0 \) ensures that

\[
L = f(x) + \sum_{i=1}^{p} \lambda_i (h_i(x) + s_i^2)
\]

has a well defined minimum.
- If \( \lambda_i < 0 \), we can always make \( L \) smaller due to the \( \lambda_is_i^2 \) term.
New first order equations

- The differentiation, $\partial L / \partial \lambda_i = 0$, ensures the constraints hold.

$$L_{\lambda_i} = 0 \Rightarrow h_i + s_i^2 = 0, \quad i = 1, \ldots, p.$$  

- The differentiation, $\frac{\partial L}{\partial s_i} = 0$, controls observed region of minimising,

$$L_{s_i} = 0 \Rightarrow 2\lambda_is_i = 0, \quad i = 1, \ldots, p.$$  

- If $\lambda_i > 0$ we are on the boundary since $s_i = 0$

- If $\lambda_i = 0$ we are inside the feasible region since $s_i \neq 0$.

- If $\lambda_i < 0$ we are on the boundary again, but the point is not a local minimum. So only focus on $\lambda_i \geq 0$. 

Simple application

Example

Consider the general problem

Minimise: \( f(x) \)

subject to: \( a \leq x \leq b \).

- Re-express the problem to allow the Kuhn-Tucker conditions.

Minimise: \( f(x) \)

subject to: \(-x \leq -a\)

\( x \leq b \).

- The Lagrangian \( L \) is given by

\[
L = f + \lambda_1(-x + s_1^2 + a) + \lambda_2(x + s_2^2 - b).
\]
Solving system

Find first order equations

\[
\frac{\partial L}{\partial x} = f'(x) - \lambda_1 + \lambda_2 = 0, \\
\frac{\partial L}{\partial \lambda_1} = -x + a + s_1^2 = 0 \\
\frac{\partial L}{\partial \lambda_2} = x - b + s_2^2 = 0, \\
\frac{\partial L}{\partial s_1} = 2\lambda_1 s_1 = 0 \\
\frac{\partial L}{\partial s_2} = 2\lambda_2 s_2 = 0.
\]

Consider the scenarios for when the constraints are at the boundary or inside the feasible region.
Possible scenarios

- **Down variable scenario (lower boundary).**

  \[ s_1 = 0 \Rightarrow \lambda_1 > 0 \text{ and } s_2^2 > 0 \Rightarrow \lambda_2 = 0. \]

  Implies that \( x = a \), and \( f'(x) = f'(a) = \lambda_1 > 0 \), so that the minimum lies at \( x = a \).
Possible scenarios

- **In variable scenario (Inside feasible region).**

\[ s_1^2 > 0 \Rightarrow \lambda_1 = 0 \text{ and } s_2^2 > 0 \Rightarrow \lambda_2 = 0. \]

Implies \( f'(x) = 0 \), for some \( x \in [a, b] \) determined by the specific \( f(x) \), and thus there is a turning point somewhere in the interval \( [a, b] \).
Possible scenarios

- **Up variable scenario** (upper boundary).

\[ s_1^2 > 0 \Rightarrow \lambda_1 = 0 \text{ and } s_2 = 0 \Rightarrow \lambda_2 > 0. \]

Implies \( x = b \), so that \( f'(x) = f'(b) = -\lambda_2 < 0 \) and we have the minimum at \( x = b \).
Numerical example

Example

Minimise: \( f(x_1, x_2, x_3) = x_1(x_1 - 10) + x_2(x_2 - 50) - 2x_3 \)
subject to:
\[ x_1 + x_2 \leq 10 \]
\[ x_3 \leq 10. \]

Write down Lagrangian,

\[ L(x, \lambda, s) = f(x) + \lambda^T(h(x) + s), \]

where,

\[ h(x) = \begin{pmatrix} x_1 + x_2 - 10 \\ x_3 - 10 \end{pmatrix} \quad s = \begin{pmatrix} s_1^2 \\ s_2^2 \end{pmatrix} \]
Solve the system of first order equations

\[
\frac{\partial L}{\partial x_1} = 2x_1 - 10 + \lambda_1 = 0 \quad (8)
\]
\[
\frac{\partial L}{\partial x_2} = 2x_2 - 50 + \lambda_1 = 0 \quad (9)
\]
\[
\frac{\partial L}{\partial x_3} = -2 + \lambda_2 = 0 \quad (10)
\]

The Kuhn-Tucker conditions are \( \lambda_1, \lambda_2 \geq 0 \) and

\[
\lambda_1(x_1 + x_2 - 10) = 0 \quad (11)
\]
\[
\lambda_2(x_3 - 10) = 0 \quad (12)
\]

Immediately from (10) and (12), \( \lambda_2 = 2 \) and \( x_3 = 10 \).
Two remaining cases to consider

- Inside the feasible set $\lambda_1 = 0$.
  
  From equations (8) and (9) find,
  
  $$\mathbf{x} = (x_1, x_2, x_3) = (5, 25, 10)$$
  
  However, constraint $x_1 + x_2 \leq 10$, is violated.
  
  Reject this solution!

- At the boundary, $\lambda_1 > 0$

  Using constraint equation (11) $x_1 + x_2 = 10$. Find the system:

  $$-2x_1 - 30 + \lambda_1 = 0$$
  $$2x_1 - 10 + \lambda_1 = 0,$$

  Solving gives, $\lambda_1 = 20$ and $x_1 = -5$, $x_2 = 15$. 
The full nonlinear optimisation problem with equality constraints

Method of Lagrange multipliers

Dealing with Inequality Constraints and the Kuhn-Tucker conditions

Second order conditions with constraints
Motivation

- Determine sufficient conditions for minima (maxima) given a set of constraints.

- Define the operator $\nabla^2$ to denote the Hessian matrices,
  \[ \nabla^2 f := H_f. \]

- What are the necessary/sufficient conditions on $\nabla^2 f$ and $\{\nabla^2 g_j\}_{j=1}^m$ that ensure a constrained minimum?

- Consider equality constraints first.

- Extend to inequality constraints.
Notation

Hessian of the constraint equations in the upcoming results needs care. Define as follows,

- Constraint Hessian involving Lagrange multipliers,

\[ \lambda^T \nabla^2 g(x) := \sum_{i=1}^{m} \lambda_i \nabla^2 g_i(x) = \sum_{i=1}^{m} \lambda_i H_{g_i}(x) \]
Necessary second order condition for a constrained minimum with equality constraints

**Theorem**

Suppose that $\mathbf{x}^*$ is a regular point of $g \in C^2$ and is a local minimum of $f$ such that $g(\mathbf{x}) = 0$ and $\mathbf{x}^*$ is a regular point of $g$. Then there exists a $\lambda \in \mathbb{R}^m$ such that,

$$\nabla f(\mathbf{x}^*) + \lambda^T \nabla g(\mathbf{x}^*) = 0.$$ 

Also, from the tangent plane $\mathcal{T} = \{\mathbf{y} : \nabla g(\mathbf{x}^*)\mathbf{y} = 0\}$, the Lagrangian Hessian at $\mathbf{x}^*$,

$$\nabla^2 L(\mathbf{x}^*) := \nabla^2 f(\mathbf{x}^*) + \lambda^T \nabla^2 g(\mathbf{x}^*),$$

is positive semi-definite on $\mathcal{T}$. i.e. $\mathbf{y}^T \nabla^2 L(\mathbf{x}^*)\mathbf{y} \geq 0$ for all $\mathbf{y} \in \mathcal{T}$. 
Sufficient second order condition for a constrained minimum with equality constraints

**Theorem**

Suppose that $x^*$ and $\lambda$ satisfy,

$$\nabla f(x^*) + \lambda^T \nabla g(x^*) = 0.$$ 

Suppose further that the Lagragian Hessian at $x^*$,

$$\nabla^2 L(x^*) := \nabla^2 f(x^*) + \lambda^T \nabla^2 g(x^*).$$

If $y^T \nabla^2 L(x^*) y > 0$ for all $y \in \mathcal{T}$. Then $f(x^*)$ is a local minimum that satisfies $g(x^*) = 0$. 

Determine if positive/negative definite

Options for determining definiteness of $\nabla^2 L(x^*)$ on the subspace $T$.

- Complete the square of $y^T \nabla^2 L(x^*) y$.
- Compute the eigenvalues of $y^T \nabla^2 L(x^*) y$.

Example

Max: $Z = x_1 x_2 + x_1 x_3 + x_2 x_3$

subject to: $x_1 + x_2 + x_3 = 3$. 
Extension: Inequality constraints

Problem:

\[
\begin{align*}
\text{Min} & \quad f(x) \\
\text{subject to:} & \quad h(x) \leq 0 \\
& \quad g(x) = 0.
\end{align*}
\]

where \( f, g, h \in C^1 \).

- Extend arguments to include both equality and inequality constraints.
- Notation: Define the index set \( \mathcal{J} \):

**Definition**

Let \( \mathcal{J}(x^*) \) be the index set of active constraints.

\[
h_j(x^*) = 0, \quad j \in \mathcal{J}; \quad h_j(x^*) < 0, \quad j \notin \mathcal{J}.
\]
First order necessary conditions

The point \( x^* \) is called a regular point of the constraints, \( g(x) = 0 \) and \( h(x) \leq 0 \), if

\[
\{ \nabla g_i \}_{i=1}^m \quad \text{and} \quad \{ \nabla h_j \}_{j \in J}
\]

are linearly independent.

Theorem

*Let \( x^* \) be a local minimum of \( f(x) \) subject the constraints \( g(x) = 0 \) and \( h(x) \leq 0 \). Suppose further that \( x^* \) is a regular point of the constraints. Then there exists \( \lambda \in \mathbb{R}^m \) and \( \mu \in \mathbb{R}^p \) with \( \mu \geq 0 \) such that,*

\[
\nabla f(x^*) + \lambda^T \nabla g(x^*) + \mu^T \nabla h(x^*) = 0
\]

\[
\mu^T h(x^*) = 0.
\]
Second order necessary conditions

**Theorem**

Let \( x^* \) be a regular point of \( g(x) = 0, \ h(x) \leq 0 \) with \( f, g, h \in C^2 \). If \( x^* \) is a local minimum point of the fully constrained problem then there exists \( \lambda \in \mathbb{R}^m \) and \( \mu \in \mathbb{R}^p \) such that (13) holds and

\[
\nabla^2 f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla^2 g_i(x^*) + \sum_{j=1}^{p} \mu_j \nabla^2 h_j(x^*)
\]

is positive semi-definite on the tangent space of the active constraints. i.e. on the space,

\[
\mathcal{T} = \left\{ d \in \mathbb{R}^n \left| \nabla g(x^*)d = 0; \ \nabla h_j(x^*)d = 0, j \in \mathcal{J} \right. \right\}
\]
Second order sufficient conditions

Theorem

Let \( x^* \) satisfy \( g(x) = 0, \ h(x) \leq 0 \). Suppose there exists \( \lambda \in \mathbb{R}^m \) and \( \mu \in \mathbb{R}^p \) such that \( \mu \geq 0 \) and (13) holds and

\[
\nabla^2 f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla^2 g_i(x^*) + \sum_{j=1}^{p} \mu_j \nabla^2 h_j(x^*)
\]

is positive definite on the tangent space of the active constraints. i.e. on the space,

\[
\mathcal{T}' = \left\{ d \in \mathbb{R}^n \middle| \nabla g(x^*)d = 0; \nabla h_j(x^*)d = 0, j \in \mathcal{J} \right\},
\]

where \( \mathcal{J} = \left\{ j \middle| h_j(x^*) = 0, \mu_j > 0 \right\} \). Then \( x^* \) is a strict local minimum of \( f \) subject to the full constraints.
Ways to compute the sufficient conditions

The nature of the Hessian of the Lagrangian function and with its auxiliary conditions can be determined.

1. Eigenvalues on the subspace $\mathcal{T}$.

2. Find the nature of the Bordered Hessian.
Eigenvalues on subspaces

Let \( \{e_i\}^m_{i=1} \) be an orthonormal basis of \( T \). Then \( d \in T \) can be written,

\[
d = E z,
\]

where \( z \in \mathbb{R}^m \) and,

\[
E = [e_1 \ e_2 \ \ldots \ e_m].
\]

Then use,

\[
d^T \nabla^2 L d = \ldots,
\]
Bordered Hessian

Alternative method to determining the behaviour of the Lagragian Hessian.

**Definition**

Given an objective function $f$ with constraints $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$, define the Bordered Hessian matrix,

$$B(x, \lambda) := \begin{pmatrix} \nabla^2 L(x, \lambda) & \nabla g(x)^T \\ \nabla g(x) & 0 \end{pmatrix}$$

Find the nature of the Bordered Hessian.