Lecture 6
Local Extremal Points and Lagrange Problems

Eivind Eriksen

BI Norwegian School of Management
Department of Economics

October 08, 2010

Local extremal points

We define local maxima and minima for functions in several variables:

> ### Definition (Local Extremal Points)
>
> *Let $f(\mathbf{x})$ be a function in n variables defined on a set $S \subseteq \mathbb{R}^n$. A point $\mathbf{x}^* \in S$ is called a local maximum point for f if*
>
> $$f(\mathbf{x}^*) \geq f(\mathbf{x}) \text{ for all } \mathbf{x} \in S \text{ close to } \mathbf{x}^*$$
>
> *and a local minimum point for f if*
>
> $$f(\mathbf{x}^*) \leq f(\mathbf{x}) \text{ for all } \mathbf{x} \in S \text{ close to } \mathbf{x}^*$$
>
> *If $\mathbf{x}^*$ is a local maximum or minimum point for f, then we call $f(\mathbf{x}^*)$ the local maximum or minimum value.*

A local extremal point is a local maximum or a local minimum point.

Global extremal points

For comparison, we review the definition of global extremal points from the previous lecture:

Definition (Global Extremal Points)

Let $f(\mathbf{x})$ be a function in n variables defined on a set $S \subseteq \mathbb{R}^n$. A point $\mathbf{x}^* \in S$ is called a global maximum point for f if

$$f(\mathbf{x}^*) \geq f(\mathbf{x}) \text{ for all } \mathbf{x} \in S$$
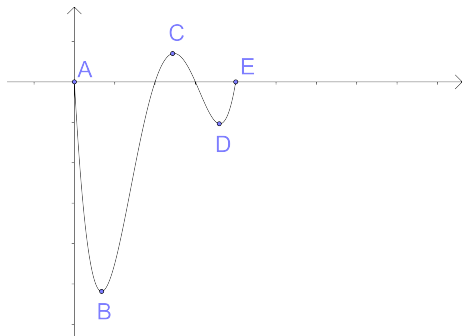
and a global minimum point for f if

$$f(\mathbf{x}^*) \leq f(\mathbf{x}) \text{ for all } \mathbf{x} \in S$$

A global extremal point is a global maximum or a global minimum point. Clearly, we have

$$\text{global extremal point} \Rightarrow \text{local extremal point}$$

Local and global extremal points: An example

We draw the graph of a function $f(x)$ in one variable defined on $[0, 4]$, and show local and global extremal points:

We see that A, B, C, D, E are local extremal points, while B and C are the only global extremal points.

First order conditions

Let $f(\mathbf{x})$ be a function in n variables defined on a set $S \subseteq \mathbb{R}^n$. We recall that a stationary point $\mathbf{x}$ for f is a solution of the first order conditions

$$f_1'(\mathbf{x}) = f_2'(\mathbf{x}) = \cdots = f_n'(\mathbf{x}) = 0$$

In general, we compute the stationary points by solving the first order conditions. This gives a system of n equations in n unknowns. In the previous example, the stationary points are B, C, D.

Proposition

If $\mathbf{x}^*$ is a local extremal point of f that is not on the boundary of S, then $\mathbf{x}^*$ is a stationary point.

Hence stationary points are candidates for local extremal points. How do we determine which stationary points are local maxima and minima?

Second order conditions

Let $f(\mathbf{x})$ be a function in n variables defined on a set $S \subseteq \mathbb{R}^n$. An interior point is a point in S that is not on the boundary of S.

> ### Theorem (Second derivative test)
>
> Let $\mathbf{x}^*$ be a interior stationary point for f. Then we have
>
> 1. If the Hessian $f''(\mathbf{x}^*)$ is positive definite, then $\mathbf{x}^*$ is a local minimum.
> 2. If the Hessian $f''(\mathbf{x}^*)$ is negative definite, then $\mathbf{x}^*$ is a local maximum.
> 3. If the Hessian $f''(\mathbf{x}^*)$ is indefinite, then $\mathbf{x}^*$ is neither a local minimum nor a local maximum.

A stationary point that is neither a local maximum nor a local minimum, is called a saddle point.

An example

Example

Let $f(x_1, x_2, x_3) = x_1^3 + 3x_1x_2 + 3x_1x_3 + x_2^3 + 3x_2x_3 + x_3^3$. Show that $\mathbf{x}^* = (-2, -2, -2)$ is a local maximum for f.

Solution

We write down the first order conditions for f, which are

$$f_1' = 3x_1^2 + 3x_2 + 3x_3 = 0$$
$$f_2' = 3x_1 + 3x_2^2 + 3x_3 = 0$$
$$f_3' = 3x_1 + 3x_2 + 3x_3^2 = 0$$

We see that $\mathbf{x}^* = (-2, -2, -2)$ is a solution. This implies that $\mathbf{x}^*$ is a stationary point. To find all stationary points would be possible, but requires more work.

An example

Solution (Continued)

Next, we find compute the Hessian of f to classify the type of the stationary point $\mathbf{x}^* = (-2, -2, -2)$:

$$f''(\mathbf{x}) = \begin{pmatrix} 6x_1 & 3 & 3 \\ 3 & 6x_2 & 3 \\ 3 & 3 & 6x_3 \end{pmatrix} \Rightarrow f''(\mathbf{x}^*) = \begin{pmatrix} -12 & 3 & 3 \\ 3 & -12 & 3 \\ 3 & 3 & -12 \end{pmatrix}$$

Since $D_1 = -12 < 0$, $D_2 = (-12)^2 - 9 = 135 > 0$ and

$$D_3 = \begin{vmatrix} -12 & 3 & 3 \\ 3 & -12 & 3 \\ 3 & 3 & -12 \end{vmatrix} = -1350 < 0$$

it follows that $\mathbf{x}^*$ is a local maximum point.

An example

Example

Let $f(x_1, x_2, x_3) = x_1^3 + 3x_1x_2 + 3x_1x_3 + x_2^3 + 3x_2x_3 + x_3^3$. Show that $\mathbf{x}^* = (0, 0, 0)$ is a saddle point for f.

Solution

We write down the first order conditions for f, which are

$$f_1' = 3x_1^2 + 3x_2 + 3x_3 = 0$$
$$f_2' = 3x_1 + 3x_2^2 + 3x_3 = 0$$
$$f_3' = 3x_1 + 3x_2 + 3x_3^2 = 0$$

We see that $\mathbf{x}^* = (0, 0, 0)$ is a solution. This implies that $\mathbf{x}^*$ is a stationary point. In fact, $(-2, -2, -2)$ and $(0, 0, 0)$ are the only stationary points.

An example

Solution (Continued)

Next, we find compute the Hessian of f to classify the type of the stationary point $\mathbf{x}^* = (0, 0, 0)$:

$$f''(\mathbf{x}) = \begin{pmatrix} 6x_1 & 3 & 3 \\ 3 & 6x_2 & 3 \\ 3 & 3 & 6x_3 \end{pmatrix} \Rightarrow f''(\mathbf{x}^*) = \begin{pmatrix} 0 & 3 & 3 \\ 3 & 0 & 3 \\ 3 & 3 & 0 \end{pmatrix}$$

Since $D_1 = 0$, the method of leading principal minors will not work in this case. Instead we find the eigenvalues:

$$\begin{vmatrix} -\lambda & 3 & 3 \\ 3 & -\lambda & 3 \\ 3 & 3 & -\lambda \end{vmatrix} = \begin{vmatrix} -\lambda & 3 & 3 \\ 3 & -\lambda & 3 \\ 6-\lambda & 6-\lambda & 6-\lambda \end{vmatrix} = (6-\lambda)(\lambda^2 + 6\lambda + 9) = 0$$

This gives eigenvalues $\lambda = 6$ and $\lambda = -3$. Hence the matrix is indefinite and $\mathbf{x}^*$ is a saddle point.

Second order condition for two variables

Let us consider the special case where f is a function in two variables ($n = 2$). In this case the Hessian of f at a stationary point $\mathbf{x}^*$ can be written as

$$f''(\mathbf{x}^*) = \begin{pmatrix} f''_{11}(\mathbf{x}^*) & f''_{12}(\mathbf{x}^*) \\ f''_{21}(\mathbf{x}^*) & f''_{22}(\mathbf{x}^*) \end{pmatrix} = \begin{pmatrix} A & B \\ B & C \end{pmatrix}$$

where $A = f''_{11}(\mathbf{x}^*), B = f''_{12}(\mathbf{x}^*)$ and $C = f''_{22}(\mathbf{x}^*)$. Since the leading principal minors are

$$D_1 = A, \quad D_2 = AC - B^2$$

we recover the second derivative test for an interior stationary point $\mathbf{x}^*$:

Second derivative test for $n = 2$

- If $AC - B^2 > 0$ and $A > 0$, then $\mathbf{x}^*$ is a local minimum.
- If $AC - B^2 > 0$ and $A < 0$, then $\mathbf{x}^*$ is a local maximum.
- If $AC - B^2 < 0$, then $\mathbf{x}^*$ is a saddle point.

Another example

Example

Let $f(x_1, x_2, x_3) = -2x_1^4 + 2x_2x_3 - x_2^2 + 8x_1$. Find all stationary points of the $f(\mathbf{x})$ and classify their type.

Solution

We write down the first order conditions for f, which are

$$f_1' = -8x_1^3 + 8 = 0$$
$$f_2' = 2x_3 - 2x_2 = 0$$
$$f_3' = 2x_2 = 0$$

From these we obtain $x_2 = 0$, $x_3 = 0$ and $x_1 = 1$, so there is one stationary point $\mathbf{x}^* = (1, 0, 0)$.

Another example

Solution (Continued)

Next, we find compute the Hessian of f to classify the type of the stationary point $\mathbf{x}^* = (1, 0, 0)$:

$$f''(\mathbf{x}) = \begin{pmatrix} -24x_1^2 & 0 & 0 \\ 0 & -2 & 2 \\ 0 & 2 & 0 \end{pmatrix} \Rightarrow f''(\mathbf{x}^*) = \begin{pmatrix} -24 & 0 & 0 \\ 0 & -2 & 2 \\ 0 & 2 & 0 \end{pmatrix}$$

We have $D_1 = -24$, $D_2 = (-24)(-2) = 48$ and $D_3 = (-24)(-4) = 96$, and it follows that the matrix is indefinite and that $\mathbf{x}^*$ is a saddle point.

Inconclusive second derivative test

In the case $n = 2$, we know that the second derivative test is inconclusive if $AC - B^2 = 0$. In this case, *all can happen* (the interior stationary point can be a local maximum, local minimum or saddle point).

Inconclusive second derivative test

If the Hessian $f''(\mathbf{x}^*)$ at an interior stationary point $\mathbf{x}^*$ is positive or negative semidefinite but not positive or negative definite (that is, if at least one of the eigenvalues is zero and all other eigenvalues have the same sign), then the second derivative test is inconclusive.

Example

Let $f_1(\mathbf{x}) = x_1^4 + x_2^4 + x_3^4$, $f_2(\mathbf{x}) = x_1^3 + x_2^3 + x_3^3$, $f_3(\mathbf{x}) = -x_1^4 - x_2^4 - x_3^4$. We see that $\mathbf{x}^* = (0, 0, 0)$ is a stationary point for f_1, f_2, f_3, and the Hessian of all three function is zero at $\mathbf{x}^* = (0, 0, 0)$. In fact, $\mathbf{x}^* = (0, 0, 0)$ is a local minimum for f_1, a saddle point for f_2 and a local maximum for f_3.

Optimization with constraints

Let $f(\mathbf{x})$ be a function defined on $S \subseteq \mathbb{R}^n$. So far, we have considered methods for finding local (and global) extremal points in the interior of S. However, we must also consider methods for finding extremal points on the boundary of S. This leads to optimization problems with constraints.

Example

Find the extremal points of $f(x_1, x_2) = x_1^2 - x_2^2$ when $x_1 \geq 0$ and $x_2 \geq 0$.

We have seen how to find extremal points in the interior of S, where $S = \{(x_1, x_2) : x_1 \geq 0, x_2 \geq 0\}$. The boundary of S is the positive x_1-axis ($x_1 \geq 0$, $x_2 = 0$) and the positive x_2-axis ($x_1 = 0$, $x_2 \geq 0$). We are therefore left with the problem of finding minimum/maximum of $f(x_1, x_2)$ when $x_1 x_2 = 0$, $x_1, x_2 \geq 0$. This is a typical optimization problem with constraints.

Lagrange problems

Lagrange problems

Let $f(\mathbf{x})$ be a function in n variables, and let

$$g_1(\mathbf{x}) = b_1, \quad g_2(\mathbf{x}) = b_2, \quad \ldots, \quad g_m(\mathbf{x}) = b_m$$

be m equality constraints, given by functions $g_1, \ldots, g_m$ and constants $b_1, \ldots, b_m$. The problem of finding the maximum or minimum of $f(\mathbf{x})$ when $\mathbf{x}$ satisfy the constraints $g_1(\mathbf{x}) = b_1, \ldots, g_m(\mathbf{x}) = b_m$ is called the Lagrange problem with objective function f and constraint functions $g_1, g_2, \ldots, g_m$.

The Lagrange problem is the general optimization problem with equality constraints. We shall describe a general method for solving Lagrange problems.

The Lagrangian

Definition

The Lagrangian or Lagrange function is the function

$$\mathcal{L}(\mathbf{x}, \lambda) = f(\mathbf{x}) - \lambda_1(g_1(\mathbf{x}) - b_1) - \lambda_2(g_2(\mathbf{x}) - b_2) - \cdots - \lambda_m(g_m(\mathbf{x}) - b_m)$$

in the $n + m$ variables $x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_m$. The variables $\lambda_1, \ldots, \lambda_m$ are called Lagrange multipliers.

To solve the Lagrange problem, we solve the system of equations consisting of the *n* first order conditions

$$\frac{\partial \mathcal{L}}{\partial x_1} = 0, \quad \frac{\partial \mathcal{L}}{\partial x_2} = 0, \quad \frac{\partial \mathcal{L}}{\partial x_3} = 0, \quad \ldots, \quad \frac{\partial \mathcal{L}}{\partial x_n} = 0$$

together with the *m* equality constraints $g_1(\mathbf{x}) = b_1, \ldots, g_m(\mathbf{x}) = b_m$. The solutions give us candidates for optimality.

The Lagrangian: An example

Example

Find the candidates for optimality for the following Lagrange problem:
Maximize/minimize the function $f(x_1, x_2) = x_1 + 3x_2$ subject to the
constraint $x_1^2 + x_2^2 = 10$.

Solution

To formulate this problem as a standard Lagrange problem, let
$g(x_1, x_2) = x_1^2 + x_2^2$ be the constraint function and let $b = 10$. (We write
g for g_1, b for b_1 and λ for λ_1 when there is only one constraint). We
form the Lagrangian

$$\mathcal{L}(x_1, x_2, \lambda) = f(\mathbf{x}) - \lambda(g(\mathbf{x}) - 10) = x_1 + 3x_2 - \lambda(x_1^2 + x_2^2 - 10)$$

The Lagrangian: An example

Solution (Continued)

The first order conditions become

$$\frac{\partial \mathcal{L}}{\partial x_1} = \frac{\partial f}{\partial x_1} - \lambda \frac{\partial g}{\partial x_1} = 1 - \lambda \cdot 2x_1 = 0 \Rightarrow x_1 = \frac{1}{2\lambda}$$

$$\frac{\partial \mathcal{L}}{\partial x_2} = \frac{\partial f}{\partial x_1} - \lambda \frac{\partial g}{\partial x_1} = 3 - \lambda \cdot 2x_2 = 0 \Rightarrow x_2 = \frac{3}{2\lambda}$$

and the constraint is $x_1^2 + x_2^2 = 10$. We may solve this system of equations, and obtain two solutions $(x_1, x_2, \lambda) = (1, 3, \frac{1}{2})$ and $(x_1, x_2, \lambda) = (-1, -3, -\frac{1}{2})$. Hence the points $(x_1, x_2) = (1, 3)$ and $(x_1, x_2) = (-1, -3)$ are the candidates for optimality.

Alternative Lagrangian

Since the first order conditions have the form

$$\frac{\partial \mathcal{L}}{\partial x_i} = 0 \quad \Leftrightarrow \quad \frac{\partial f}{\partial x_i} - \lambda_1 \frac{\partial g_1}{\partial x_i} - \cdots - \lambda_m \frac{\partial g_m}{\partial x_i} = 0$$

for $1 \leq i \leq n$, they are unaffected by the constants b_i. It is therefore possible to use the following alternative Lagrangian:

$$\mathcal{L}(\mathbf{x}, \lambda) = f(\mathbf{x}) - \lambda_1 g_1(\mathbf{x}) - \lambda_2 g_2(\mathbf{x}) - \cdots - \lambda_m g_m(\mathbf{x})$$

This alternative Lagrangian is simpler. The advantage of the first form of Lagrangian is that the constraints can be obtained as

$$\frac{\partial \mathcal{L}}{\partial \lambda_i} = 0 \quad \Leftrightarrow \quad g_i(\mathbf{x}) = b_i$$

Non-degenerate constraint qualification

The NDCQ condition

The non-degenerate constraint qualification (or NDCQ) condition is satisfied if

$$
\mathrm{rk}\begin{pmatrix}
\frac{\partial g_1}{\partial x_1}(\mathbf{x}) & \frac{\partial g_1}{\partial x_2}(\mathbf{x}) & \ldots & \frac{\partial g_1}{\partial x_n}(\mathbf{x}) \\
\frac{\partial g_2}{\partial x_1}(\mathbf{x}) & \frac{\partial g_2}{\partial x_2}(\mathbf{x}) & \ldots & \frac{\partial g_2}{\partial x_n}(\mathbf{x}) \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial g_m}{\partial x_1}(\mathbf{x}) & \frac{\partial g_m}{\partial x_2}(\mathbf{x}) & \ldots & \frac{\partial g_m}{\partial x_n}(\mathbf{x})
\end{pmatrix} = m
$$

Note that $m < n$ in most Lagrange problems (that is, the number of constraints is less than the number of variables). If not, the set of points satisfying the constraints would usually not allow for any degrees of freedom.

Necessary conditions

Theorem

Let the functions $f, g_1, \ldots, g_m$ in a Lagrange problem be defined on a subset $S \subseteq \mathbb{R}^n$. If $\mathbf{x}^*$ is a point in the interior of S that solves the Lagrange problem and satisfy the NDCQ condition, then there exist unique numbers $\lambda_1, \ldots, \lambda_m$ such that $x_1^*, \ldots, x_n^*, \lambda_1, \ldots, \lambda_m$ satisfy the first order conditions.

Important remark

This theorem implies that any solution of the Lagrange problem can be extended to a solution of the first order conditions. This means that if we solve the system of equations that includes the first order conditions and the constraints, we obtain candidates for optimality.

The Lagrangian: An example

Example

Maximize/minimize the function $f(x_1, x_2) = x_1 + 3x_2$ subject to the constraint $x_1^2 + x_2^2 = 10$.

Solution

We found candidates for optimality by solving the first order conditions and the constraint earlier, and found two solutions $(1, 3)$ and $(-1, -3)$. We compute $f(1, 3) = 10$ and $f(-1, -3) = -10$. Also, the NDCQ is satisfied in those points, since $\begin{pmatrix} g'_x & g'_y \end{pmatrix} = \begin{pmatrix} 2x & 2y \end{pmatrix}$ is non-zero at $(1, 3)$ and $(-1, -3)$. We conclude that if there is a maximum point, then it must be $(1, 3)$, and if there is a minimum point, then it must be $(-1, -3)$.

Extreme value theorem

- A set $S \subseteq \mathbb{R}^n$ is closed if it contains its boundary. Subsets defined by equalities and inequalities defined by $\geq$ or $\leq$ are usually closed.
- A set $S \subseteq \mathbb{R}^n$ is bounded if all points in S are contained in a ball with a finite radius. For instance, $\{(x, y) : 0 \leq x \leq 1, \ 0 \leq y \leq 2\}$ is bounded, while the first quadrant $\{(x, y) : x \geq 0, \ y \geq 0\}$ is not.

Theorem

Let $f(\mathbf{x})$ be a continuous function defined on a closed and bounded subset $S \subseteq \mathbb{R}^n$. Then f has a maximum point and a minimum point in S.

Sufficient conditions

Theorem

Let $\mathbf{x}^*$ be a point satisfying the constraints, and assume that there exist numbers $\lambda_1^*, \ldots, \lambda_m^*$ such that $(x_1^*, \ldots, x_n^*, \lambda_1^*, \ldots, \lambda_m^*)$ satisfy the first order conditions. Then we have:

1. *If $\mathcal{L}(\mathbf{x})$ is convex as a function in $\mathbf{x}$ (with $\lambda_i = \lambda_i^*$ fixed), then $\mathbf{x}^*$ is a solution to the minimum Lagrange problem.*

2. *If $\mathcal{L}(\mathbf{x})$ is concave as a function in $\mathbf{x}$ (with $\lambda_i = \lambda_i^*$ fixed), then $\mathbf{x}^*$ is a solution to the maximum Lagrange problem.*

Let us review the previous example, using this theorem:

Example

Maximize/minimize the function $f(x_1, x_2) = x_1 + 3x_2$ subject to the constraint $x_1^2 + x_2^2 = 10$.

Sufficient conditions: An example

Solution

The Lagrangian for this problem is

$$\mathcal{L}(x, y, \lambda) = x + 3y - \lambda(x^2 + y^2 - 10)$$

We found one solution $(x, y, \lambda) = (1, 3, \frac{1}{2})$ of the first order conditions, and when we fix $\lambda = \frac{1}{2}$, we get the Lagrangian

$$\mathcal{L}(x, y) = x + 3y - \frac{1}{2}(x^2 + y^2) - 5$$

This function is clearly concave, so $(x, y) = (1, 3)$ is a maximum point. Similarly, $(x, y) = (-1, -3)$ is a minimum point since $\mathcal{L}(x, y)$ is convex when we fix $\lambda = -\frac{1}{2}$.

Sufficient conditions: Another example

Example

Solve the first order conditions for the following Lagrange problem:
Maximize $f(x_1, x_2, x_3) = x_1 x_2 x_3$ subject to $g_1(x_1, x_2, x_3) = x_1^2 + x_2^2 = 1$
and $g_2(x_1, x_2, x_3) = x_1 + x_3 = 1$.

Solution

The NDCQ condition is that the following matrix has rank two:

$$\begin{pmatrix} \frac{\partial g_1}{\partial x_1}(\mathbf{x}) & \frac{\partial g_1}{\partial x_2}(\mathbf{x}) & \frac{\partial g_1}{\partial x_3}(\mathbf{x}) \\ \frac{\partial g_2}{\partial x_1}(\mathbf{x}) & \frac{\partial g_2}{\partial x_2}(\mathbf{x}) & \frac{\partial g_2}{\partial x_3}(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} 2x_1 & 2x_2 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

This is clearly the case unless $x_1 = x_2 = 0$, and this is impossible because of the first constraint. So the NDCQ is satisfied for all points satisfying the constraints.

Sufficient conditions: Another example

Solution (Continued)

The first order conditions are given by

$$\frac{\partial \mathcal{L}}{\partial x_1} = x_2 x_3 - \lambda_1 \cdot 2x_1 - \lambda_2 \cdot 1 = 0$$

$$\frac{\partial \mathcal{L}}{\partial x_2} = x_1 x_3 - \lambda_1 \cdot 2x_2 - \lambda_2 \cdot 0 = 0$$

$$\frac{\partial \mathcal{L}}{\partial x_3} = x_1 x_2 - \lambda_1 \cdot 0 - \lambda_2 \cdot 1 = 0$$

We solve the last two equations for λ_1 and λ_2, and get

$$\lambda_1 = \frac{x_1 x_3}{2x_2}, \quad \lambda_2 = x_1 x_2$$

We substitute this into the first equation, and get

Sufficient conditions: Another example

Solution (Continued)

$$x_2 x_3 - \frac{x_1 x_3}{2x_2} 2x_1 - x_1 x_2 = 0 \quad \Leftrightarrow \quad x_2^2 x_3 - x_1^2 x_3 - x_1 x_2^2 = 0$$

Finally, we use that $x_2^2 = 1 - x_1^2$ and $x_3 = 1 - x_1$ from the constraints, and get the equation

$$(1 - x_1^2)(1 - x_1) - x_1^2(1 - x_1) - x_1(1 - x_1^2) = 0$$

We see that $(1 - x_1)$ is a common factor, and obtain the equation

$$(1 - x_1)(1 - x_1^2 - x_1^2 - x_1(1 + x_1)) = (1 - x_1)(-3x_1^2 - x_1 + 1) = 0$$

This gives $x_1 = 1$ and $x_1 = \frac{1}{6}(-1 \pm \sqrt{13})$. Solving for $x_2, x_3, \lambda_1, \lambda_2$, we get 5 solutions in total (see the lecture notes for details).