Lecture 3

Lagrange methods

Lagrange methods

The Lagrange method is a very important method for optimizing constrained functions. The same method is used in the support vector machine.

In mathematical optimization method of Lagrange multipliers is a strategy for finding the local maxima and minima function subject to equation constraints (i.e., subject to the condition that one or more equations have to be satisfied exactly by the chosen values of the variables).

Lagrange methods

The basic idea is to convert a constrained problem into a form such that the derivative test of an unconstrained problem can still be applied.

In the general case, the Lagrangian is defined as:

$$L(x, \lambda) = f(x) + \langle \lambda, g(x) \rangle$$

where λ – the Lagrange multiplier.

In the simple case, this simplifies to

$$L(x, \lambda) = f(x) + \lambda g(x)$$

Lagrange methods

So, in order to find the maximum or minimum of a function $f(x)$ subjected to the equality constraint $g(x) = 0$ find the stationary points of $L(x, \lambda)$. This means that all partial derivatives should be zero, including the partial derivative with respect to λ:

$$\frac{\partial L}{\partial x} = 0 \text{ and } \frac{\partial L}{\partial \lambda} = 0$$

or equivalently $\frac{\partial f(x)}{\partial x} + \lambda \frac{\partial g(x)}{\partial x} = 0$ and $g(x) = 0$.

Lagrange methods

The solution corresponding to the original constrained optimization is always a saddle point of the Lagrangian function, which can be identified among the stationary points from the definiteness of the bordered Hessian matrix.

Lagrange methods. Single constraint

For the case of only one constraint and only two choice variables, consider the optimization problem:

$$f(x,y) \to max$$
$$g(x,y) = 0$$

We assume that both f and g have continuous first partial derivatives.

Lagrange function looks like:

$$L(x,y,\lambda) = f(x,y) + \lambda g(x,y)$$

Lagrange methods. Single constraint

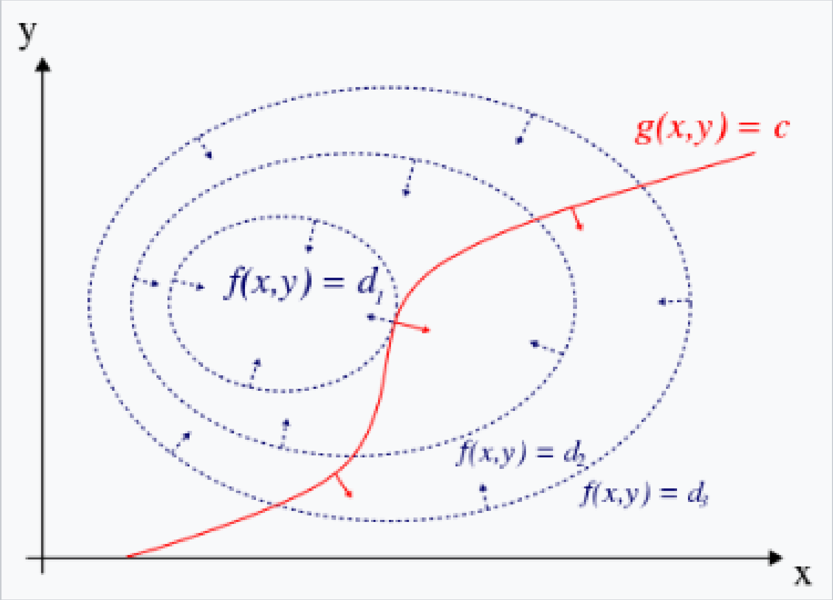


Figure 1: The red curve shows the constraint $g(x, y) = c$. The blue curves are contours of $f(x, y)$. The point where the red constraint tangentially touches a blue contour is the maximum of $f(x, y)$ along the constraint, since $d_1 > d_2$.

Lagrange methods. Single constraint

If $f(x_0, y_0)$ is a maximum of $f(x, y)$ for the original constrained problem and $\nabla g(x_0, y_0) \neq 0$ then there exists λ_0 such that (x_0, y_0, λ_0) is a stationary point for the Lagrange function.

Suppose we walk along the contour line with $g = c$. We are interested in finding points where f almost does not change as we walk, since these points might be maxima.

Lagrange methods. Single constraint

There are two ways this could happen:

We could touch a contour line of *f*, since by definition *f* does not change as we walk along its contour lines. This would mean that the tangents to the contour lines of f and *g* are parallel here.

We have reached a "level" part of *f*, meaning that *f* does not change in any direction.

Lagrange methods. Single constraint

To check the first possibility, notice that since the gradient of a function is perpendicular to the contour lines, the tangents to the contour lines of f and g are parallel if and only if the gradients of f and g are parallel.

Thus we want points (x, y) where $\mathrm{g}(x, y) = c$ and $\nabla_{x,y} f = \lambda \nabla_{x,y} g$

where $\nabla_{x,y} f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right)$, $\nabla_{x,y} \mathrm{g} = \left(\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}\right)$ are the respective gradients.

Lagrange methods. Single constraint

To incorporate these conditions into one equation, we introduce an auxiliary function:

$$L(x, y, \lambda) = f(x, y) + \lambda g(x, y)$$

and solve:

$$\nabla_{x,y,\lambda} L(x, y, \lambda) = 0$$

Note that $\nabla_\lambda L(x, y, \lambda) = 0$ implies $g(x, y) = 0$.

So, $\nabla_{x,y,\lambda} L(x, y, \lambda) = 0 \Leftrightarrow \begin{cases} \nabla_{x,y} f(x, y) = -\lambda \nabla_{x,y} \mathrm{g}(x, y) \\ \mathrm{g}(x, y) = 0 \end{cases}$

Lagrange methods. Algorithm

Algorithm for solving a nonlinear programming problem with equality constraints using the Lagrange multiplier method:

1. Compose the Lagrange function $L(x, \lambda_0, \dots, \lambda_m) = \lambda_0 f(x) + \sum_{i=1}^{m} g_i(x)\lambda_i$
2. Write down the necessary conditions for an extremum:

$$\begin{cases} \lambda_0 \dfrac{\partial f(x)}{\partial x_j} + \lambda_i \displaystyle\sum_{i=1}^{m} \dfrac{\partial g_i(x)}{\partial x_j} = 0, j = 1, \dots, n \\ g_i(x) = 0, i = 1, \dots, m \end{cases}$$

Lagrange methods. Algorithm

3. Solve the resulting system. If x is a solution of task

$$\begin{cases} f(x) \to \max(\min) \\ g_i(x) = 0, i = 1, \dots, m \end{cases}$$

then, according to the theorem on the necessary condition for an extremum, there exists a vector $\lambda \neq 0$ such that a couple (x, λ) satisfies the system under consideration. This means that we will look for the extreme points of the problem under consideration among the solutions of the system for which $\lambda \neq 0$.

Lagrange methods. Algorithm

4. We divide the solutions of system into two groups: irregular $(\lambda_0 = 0)$ and regular $(\lambda_0 \neq 0)$. For regular solutions we assume that $\lambda_0 = C\ (for\ example\ C = 1)$.

5. Check whether the found points are extremal solutions to our problem. This can be done either by Weierstrass's theorem or by using the theorem on the sufficient condition of an extremum (in the case of twice continuously differentiable functions $f(x)$ и $g_i(x), i = 1, \dots, m$).

Lagrange methods. Algorithm

Theorem (First-Order Necessary Condition for a Local Extremum in Problems with Equality Constraints) Let $f(x), g_i(x), i = 1, \ldots, m$ be functions continuously differentiable in a neighborhood of the point x^*. If the point x^* is a local extremum of the constrained optimization problem with equality constraints , then there exists a non-zero set of Lagrange multipliers λ such that the stationarity condition with respect to x holds for the Lagrange function:

$$L'_x(x^*, \lambda) = 0$$

Next, we formulate the sufficient condition for an extremum.

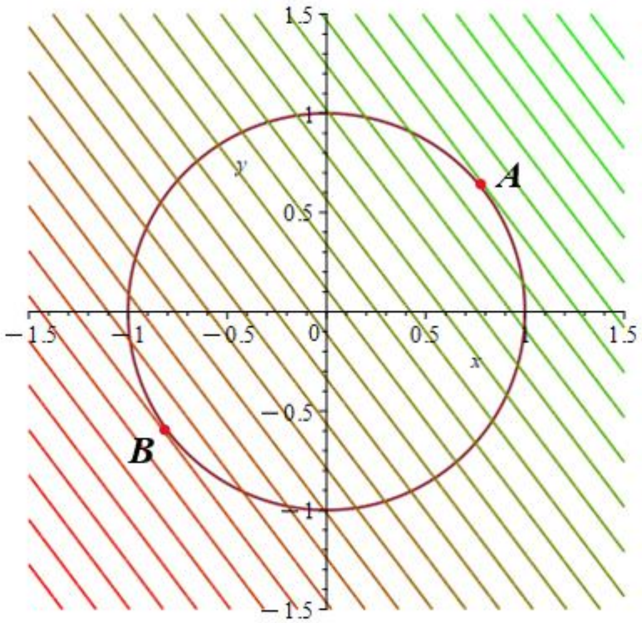
Lagrange methods. Algorithm

Theorem (First-Order Sufficient Condition for a Local Extremum in Problems with Equality Constraints) Let $x \in \Omega$ and suppose there exists a vector λ such that $\lambda_0 \geq 0$ and $L'_x(x, \lambda) = 0$. If for any non-zero vector dx satisfying the conditions $g'_i(x)dx = 0, i = 1, ..., m$, the inequality $L''_{xx}(x, \lambda)dx^2 > 0$ holds, then the function $f(x)$ attains a strict local minimum at the point x. If, however, $L''_{xx}(x, \lambda)dx^2 < 0$, then the function $f(x)$ attains a strict local maximum at the point x.

Lagrange methods. Example

Example. Solve the constrained optimization problem:

$$\begin{cases} f(x,y) = 4x + 3y \to \max(\min) \\ x^2 + y^2 - 1 = 0 \end{cases}$$



Lagrange methods. Example

Let us compose the Lagrange function of the problem under consideration:

$$L(x, y, \lambda_0, \lambda_1) = \lambda_0(4x + 3y) + \lambda_1(x^2 + y^2 - 1)$$

Let us find the partial derivatives of the Lagrange function with respect to all variables:

$$\frac{\partial L(x, y, \lambda)}{\partial x} = 4\lambda_0 + 2\lambda_1 x; \; \frac{\partial L(x, y, \lambda)}{\partial y} = 3\lambda_0 + 2\lambda_1 y$$

Lagrange methods. Example

The necessary condition for an extremum is:

$$\begin{cases} \dfrac{\partial L(x, y, \lambda)}{\partial x} = 0 \\ \dfrac{\partial L(x, y, \lambda)}{\partial y} = 0 \\ x^2 + y^2 - 1 = 0 \end{cases}$$

$$\begin{cases} 4\lambda_0 + 2\lambda_1 x = 0 \\ 3\lambda_0 + 2\lambda_1 y = 0 \\ x^2 + y^2 - 1 = 0 \end{cases}$$

Lagrange methods. Example

Let's look for irregular solutions to the system. To do this, let's set $\lambda_0 = 0$ and get the following system:

$$\begin{cases} 2\lambda_1 x = 0 \\ 2\lambda_1 y = 0 \\ x^2 + y^2 - 1 = 0 \end{cases}$$

The system has a solution at $\lambda_1 = 0$, i.e. $\lambda = (0,0)$, i.e. there are no points suspicious of an extremum.

Let's move on to finding regular solutions to the system. We assume $\lambda_0 = 1$:

Lagrange methods. Example

$$\begin{cases} 4 + 2\lambda_1 x = 0 \\ 3 + 2\lambda_1 y = 0 \\ x^2 + y^2 - 1 = 0 \end{cases}$$

Solution of this system $\lambda_1 = \pm\frac{5}{2}$ leads to two points that are suspected of being extreme: $A\left(\frac{4}{5};\frac{3}{5}\right);\ B\left(-\frac{4}{5};-\frac{3}{5}\right)$.

In this example, it is convenient to use the Weierstrass theorem. It is applicable because we are looking for the extremum on the circle, which is a compact set, and the objective function is continuous. This means that f(x,y) reaches its maximum and minimum on the circle.

```python
import sympy as sp
def obj_func(x, y):
    return 4 * x + 3 * y

def constr(x, y):
    return x**2+y**2-1



x, y, l0, l1 = sp.symbols('x y l0 l1')

lagrange = l0 * (4 * x + 3 * y) + l1 * (x**2 + y**2 - 1)



df_dx = sp.diff(lagrange, x)
df_dy = sp.diff(lagrange, y)
df_dl0 = sp.diff(lagrange, l0)
df_dl1 = sp.diff(lagrange, l1)

print("Partial derivative with respect to x:", df_dx)
print("Partial derivative with respect to y:", df_dy)
print("Partial derivative with respect to l0:", df_dl0)
print("Partial derivative with respect to l1:", df_dl1)
```

```
Partial derivative with respect to x: 4*l0 + 2*l1*x
Partial derivative with respect to y: 3*l0 + 2*l1*y
Partial derivative with respect to l0: 4*x + 3*y
Partial derivative with respect to l1: x**2 + y**2 - 1
```

```
solutions=sp.solve([df_dx,df_dy,df_dl1],(x,y,l1),dict=True)

for sol in solutions:
  print(sol)
```

```
{l1: -5*10/2, x: 4/5, y: 3/5}
{l1: 5*10/2, x: -4/5, y: -3/5}
```

Lagrange methods. Example

Example. Solve the constrained optimization problem:

$$\begin{cases} f(x) = 7x_1x_2 + x_2x_3 \to \max(\min) \\ 3x_1 + x_2 = 21 \\ x_2 + x_3 = 10 \end{cases}$$

Lagrange methods. Example

Let us compose the Lagrange function of the problem under consideration:

$$L(x,\lambda) = \lambda_0(7x_1x_2 + x_2x_3) + \lambda_1(3x_1 + x_2 - 21) + \lambda_2(x_2 + x_3 - 10)$$

Let us find the partial derivatives of the Lagrange function with respect to all variables:

$$\frac{\partial L(x,\lambda)}{\partial x_1} = 7\lambda_0 x_2 + 3\lambda_1; \ \frac{\partial L(x,\lambda)}{\partial x_2} = 7\lambda_0 x_1 + \lambda_0 x_3 + \lambda_1 + \lambda_2;$$

$$\frac{\partial L(x,\lambda)}{\partial x_3} = \lambda_0 x_2 + \lambda_2$$

Lagrange methods. Example

The necessary condition for an extremum is:

$$\begin{cases} 7\lambda_0 x_2 + 3\lambda_1 = 0 \\ 7\lambda_0 x_1 + \lambda_0 x_3 + \lambda_1 + \lambda_2 = 0 \\ \lambda_0 x_2 + \lambda_2 = 0 \\ 3x_1 + x_2 = 21 \\ x_2 + x_3 = 10 \end{cases}$$

Let's look for irregular solutions to the system. To do this, we set $\lambda_0 = 0$ and obtain the following system:

Lagrange methods. Example

$$\begin{cases} 3\lambda_1 = 0 \\ \lambda_1 + \lambda_2 = 0 \\ \lambda_2 = 0 \\ 3x_1 + x_2 = 21 \\ x_2 + x_3 = 10 \end{cases}$$

The system has a solution at $\lambda_1, \lambda_2 = 0$, i.e. $\lambda = (0,0,0)$, i.e. there are no points suspicious of an extremum. Let's move on to finding regular solutions to the system. We assume $\lambda_0 = 1$:

Lagrange methods. Example

$$\begin{cases} 7x_2 + 3\lambda_1 = 0 \\ 7x_1 + x_3 + \lambda_1 + \lambda_2 = 0 \\ x_2 + \lambda_2 = 0 \\ 3x_1 + x_2 = 21 \\ x_2 + x_3 = 10 \end{cases}$$

The solution of this system $\lambda_1 = -20\frac{13}{20};\ \lambda_2 = -8\frac{17}{20}$ leads to a point that is suspected of being an extremum: $A\left(4\frac{1}{20}; 8\frac{17}{20}; 1\frac{3}{20}\right)$.

Let us use the theorem on the sufficient condition for an extremum to check whether point A is a local minimum or local maximum.

Lagrange methods. Example

From the condition $g_i'(x)dx = 0, i = 1, \dots, m$ of the theorem follows the system of equations:

$$\begin{cases} d(3x_1 + x_2 - 21) = 0 \\ d(x_2 + x_3 - 10) = 0 \end{cases}$$

Where from the properties of the differential:

$$\begin{cases} 3dx_1 + dx_2 = 0 \\ dx_2 + dx_3 = 0 \end{cases} \Rightarrow \begin{cases} dx_1 = -\frac{1}{3}dx_2 \\ dx_3 = -dx_2 \end{cases}$$

Lagrange methods. Example

Next, we define all the partial derivatives of the second order of the Laplace function:

$$\frac{\partial^2 L(x,\lambda)}{\partial x_1^2} = 0;\ \frac{\partial^2 L(x,\lambda)}{\partial x_2^2} = 0; \frac{\partial^2 L(x,\lambda)}{\partial x_3^2} = 0;\ \frac{\partial^2 L(x,\lambda)}{\partial x_1 \partial x_2} = \frac{\partial^2 L(x,\lambda)}{\partial x_2 \partial x_1} = 7;$$

$$\frac{\partial^2 L(x,\lambda)}{\partial x_1 \partial x_3} = \frac{\partial^2 L(x,\lambda)}{\partial x_3 \partial x_1} = 0;\ \frac{\partial^2 L(x,\lambda)}{\partial x_2 \partial x_3} = \frac{\partial^2 L(x,\lambda)}{\partial x_3 \partial x_2} = 1$$

Let us form the second differential of the Laplace function. To do this, let us recall that the second differential of a function of three variables has the form:

Lagrange methods. Example

$$d^2f = \frac{\partial^2 f(x,y,z)}{\partial x^2}dx^2 + \frac{\partial^2 f(x,y,z)}{\partial y^2}dy^2 + \frac{\partial^2 f(x,y,z)}{\partial z^2}dz^2 + 2\frac{\partial^2 f(x,y,z)}{\partial x \partial y}dxdy + 2\frac{\partial^2 f(x,y,z)}{\partial x \partial z}dxdz + 2\frac{\partial^2 f(x,y,z)}{\partial y \partial z}dydz$$

The second differential of the Laplace function at point A has the form:

$$d^2L(A,\lambda) = 0dx_1^2 + 0dx_2^2 + 0dx_3^2 + 14dx_1dx_2 + 0dx_1dx_3 + 2dx_2dx_3 = 14dx_1dx_2 + 2dx_2dx_3$$

Lagrange methods. Example

Let's use the expression obtained earlier $dx_1 = -\frac{1}{3}dx_2,\ dx_3 = -dx_2$ and substitute this expression into the second differential of the Laplace function:

$$d^2L(A,\lambda) = 14dx_1dx_2 + 2dx_2dx_3 = -\frac{14}{3}(dx_2)^2 - 2(dx_2)^2 = -\frac{20}{3}(dx_2)^2 < 0$$

Since the second differential of the Laplace function at point A is less than zero, this means that point A is a strict local maximum.

```python
import sympy as sp
import numpy as np
from fractions import Fraction

# Define symbolic variables
x1, x2, x3 = sp.symbols('x1 x2 x3', real=True)
l0, l1, l2 = sp.symbols('λ0 λ1 λ2', real=True)
dx1, dx2, dx3 = sp.symbols('dx1 dx2 dx3', real=True)



# Step 1: Define objective function and constraints
f = 7*x1*x2 + x2*x3
g1 = 3*x1 + x2 - 21
g2 = x2 + x3 - 10

print("\n1. PROBLEM FORMULATION")
print(f"f(x) = {f}")
print(f"g1 = {g1} = 0")
print(f"g2 = {g2} = 0")

# Step 2: Lagrange function
L = l0*f + l1*g1 + l2*g2
print(f"\n2. LAGRANGE FUNCTION")
print(f"L = {L}")
```

```python
# Step 3: Partial derivatives
dL_dx1 = sp.diff(L, x1)
dL_dx2 = sp.diff(L, x2)
dL_dx3 = sp.diff(L, x3)

print("\n3. PARTIAL DERIVATIVES")
print(f"∂L/∂x1 = {dL_dx1} = 0")
print(f"∂L/∂x2 = {dL_dx2} = 0")
print(f"∂L/∂x3 = {dL_dx3} = 0")

# Step 4: Case λ0 = 0 (irregular)
print("\n4. CASE λ0 = 0 (IRREGULAR)")
dL_dx1_0 = dL_dx1.subs(l0, 0)
dL_dx2_0 = dL_dx2.subs(l0, 0)
dL_dx3_0 = dL_dx3.subs(l0, 0)

print(f"∂L/∂x1 = {dL_dx1_0} = 0 → 3λ1 = 0 → λ1 = 0")
print(f"∂L/∂x2 = {dL_dx2_0} = 0 → λ1 + λ2 = 0 → λ2 = 0")
print(f"∂L/∂x3 = {dL_dx3_0} = 0 → λ2 = 0")
print(f"Result: λ = (0, 0, 0) - no suspicious points")
```

```python
# Step 5: Case λ0 = 1 (regular)
print("\n5. CASE λ0 = 1 (REGULAR)")
eq1 = 7*x2 + 3*l1
eq2 = 7*x1 + x3 + l1 + l2
eq3 = x2 + l2
eq4 = 3*x1 + x2 - 21
eq5 = x2 + x3 - 10

system = [eq1, eq2, eq3, eq4, eq5]
vars = [x1, x2, x3, l1, l2]

sol = sp.solve(system, vars, dict=True)[0]

x1_val = float(sol[x1])
x2_val = float(sol[x2])
x3_val = float(sol[x3])
l1_val = float(sol[l1])
l2_val = float(sol[l2])

print(f"Solution:")
print(f"λ1 = {l1_val:.6f} = -20 13/20")
print(f"λ2 = {l2_val:.6f} = -8 17/20")
print(f"Point A: x1 = {x1_val:.6f} = 4 1/20")
print(f"         x2 = {x2_val:.6f} = 8 17/20")
print(f"         x3 = {x3_val:.6f} = 1 3/20")
```

```python
# Step 6: Constraint differentials
print("\n6. CONSTRAINT DIFFERENTIALS")
print(f"d(g1) = 3dx1 + dx2 = 0 → dx1 = -dx2/3")
print(f"d(g2) = dx2 + dx3 = 0 → dx3 = -dx2")

# Step 7: Second differential
print("\n7. SECOND DIFFERENTIAL")
d2L = 14*dx1*dx2 + 2*dx2*dx3
print(f"d²L = {d2L}")

d2L_sub = d2L.subs({dx1: -dx2/3, dx3: -dx2})
d2L_sub = sp.simplify(d2L_sub)
print(f"After substitution: d²L = -20/3·(dx2)²")

# Step 8: Sign analysis
print("\n8. SIGN ANALYSIS")
print(f"d²L = -20/3·(dx2)²")
print(f"Coefficient: -20/3 < 0")
print(f"(dx2)² > 0 for any dx2 ≠ 0")
print(f"Therefore d²L < 0 for all admissible directions")

# Step 9: Final conclusion
print("\n9. CONCLUSION")
print("=" * 60)
print(f"Point A({x1_val:.3f}; {x2_val:.3f}; {x3_val:.3f}) is a")
print("STRICT LOCAL MAXIMUM of f(x) subject to the constraints.")
print("=" * 60)
```

```
1. PROBLEM FORMULATION
f(x) = 7*x1*x2 + x2*x3
g1 = 3*x1 + x2 - 21 = 0
g2 = x2 + x3 - 10 = 0

2. LAGRANGE FUNCTION
L = λ0*(7*x1*x2 + x2*x3) + λ1*(3*x1 + x2 - 21) + λ2*(x2 + x3 - 10)

3. PARTIAL DERIVATIVES
∂L/∂x1 = 7*x2*λ0 + 3*λ1 = 0
∂L/∂x2 = λ0*(7*x1 + x3) + λ1 + λ2 = 0
∂L/∂x3 = x2*λ0 + λ2 = 0

4. CASE λ0 = 0 (IRREGULAR)
∂L/∂x1 = 3*λ1 = 0 → 3λ1 = 0 → λ1 = 0
∂L/∂x2 = λ1 + λ2 = 0 → λ1 + λ2 = 0 → λ2 = 0
∂L/∂x3 = λ2 = 0 → λ2 = 0
Result: λ = (0, 0, 0) - no suspicious points

5. CASE λ0 = 1 (REGULAR)
Solution:
λ1 = -20.650000 = -20 13/20
λ2 = -8.850000 = -8 17/20
Point A: x1 = 4.050000 = 4 1/20
         x2 = 8.850000 = 8 17/20
         x3 = 1.150000 = 1 3/20

6. CONSTRAINT DIFFERENTIALS
d(g1) = 3dx1 + dx2 = 0 → dx1 = -dx2/3
d(g2) = dx2 + dx3 = 0 → dx3 = -dx2
```

```
7. SECOND DIFFERENTIAL
d²L = 14*dx1*dx2 + 2*dx2*dx3
After substitution: d²L = -20/3·(dx2)²

8. SIGN ANALYSIS
d²L = -20/3·(dx2)²
Coefficient: -20/3 < 0
(dx2)² > 0 for any dx2 ≠ 0
Therefore d²L < 0 for all admissible directions

9. CONCLUSION
============================================================
Point A(4.050; 8.850; 1.150) is a
STRICT LOCAL MAXIMUM of f(x) subject to the constraints.
============================================================
```

```python
# Step 10: Verification
f_val = 7*x1_val*x2_val + x2_val*x3_val
print(f"\nFunction value at maximum: f(A) = {f_val:.6f}")
```