MATH20070/2020 - Optimization in Finance

2021/2022

Contents

Chapter 1

Unconstrained Optimization

1.1 Introduction

[1] We derive a Second Derivative Test for function of n-variables. In order to do this we will combine techniques from calculus of several variables and linear algebra.

1.2 Background and Notation

We let $\mathbb{R}$ denote the set of real numbers.

For $a, b \in \mathbb{R}$ with $a \leq b$, we let

$$[a, b] := \{x \in \mathbb{R} : a \leq x \leq b\}$$

be the closed interval with endpoints a and b and

$$(a, b) := \{x \in \mathbb{R} : a < x < b\}$$

be the open interval with endpoints a and b.

Let us recall how to find and classify extreme points of functions of one variable. Let $f\colon (a, b) \to \mathbb{R}$ be a differentiable function. We say that $c \in (a, b)$ is a critical point of f if $f'(c) = 0$. The follows Theorem tells us if we have a local maximum, a local minimum or a point of inflection.

Theorem 1.1 (Second Derivative Test) Let $f\colon (a, b) \to \mathbb{R}$ be a twice differentiable function and $c \in (a, b)$ be a critical point of f (that is, $f'(c) = 0$).

(i) If $f''(c) > 0$, then f has a relative minimum at c.

(ii) If $f''(c) < 0$, then f has a relative maximum at c.

(iii) If $f''(c) = 0$ AND f'' changes sign when passing through c then c is an inflection point for f

1.2.1 Partial Derivatives

Let us start by considering what happens for functions of two variables. Let $f(x, y)$ be a function of two variables x and y. Then $f(x, y)$ has two first order partial derivatives:

$$f_x = \frac{\partial f}{\partial x}$$

is the derivative with respect to x, obtained by differentiating f with respect to x treating y as a constant,

$$f_y = \frac{\partial f}{\partial y}$$

is the derivative with respect to y, obtained by differentiating f with respect to y treating x as a constant.

Formally we obtain $\dfrac{\partial f}{\partial x}$ and $\dfrac{\partial f}{\partial y}$ by the formulae

$$\frac{\partial f}{\partial x}(x, y) = \lim_{h \to 0} \frac{f(x + h, y) - f(x, y)}{h}$$

and

$$\frac{\partial f}{\partial y}(x, y) = \lim_{k \to 0} \frac{f(x, y + k) - f(x, y)}{k}.$$

We call f_x and f_y the first order partial derivatives of f with respect to x and y respectively.

Example 1.2 If $f(x, y) = xe^{x^2+y^2}$ then

$$\frac{\partial f}{\partial x}(x, y) = (2x^2 + 1)e^{x^2+y^2},$$

and

$$\frac{\partial f}{\partial y}(x, y) = 2xye^{x^2+y^2}.$$

For functions of n variables, $f(x_1, x_2, \ldots, x_n)$ we have n first order partial derivatives

$$f_{x_1} = \frac{\partial f}{\partial x_1}, \qquad f_{x_2} = \frac{\partial f}{\partial x_2}, \ldots, f_{x_n} = \frac{\partial f}{\partial x_n}.$$

The partial derivative f_{x_j} is obtained by differentiating f with respect to x_j treating the other $n - 1$ variables $x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n$ as constants.

Example 1.3 If

$$f(x, y, z) = 5x^2 + 4xy + y^3 - 2y + z^5 + xe^z$$

then

$$f_x(x, y, z) = 10x + 4y + e^z, \qquad f_y(x, y, z) = 4x + 3y^2 - 2,$$

$$f_z(x, y, z) = 5y^4 + xe^z.$$

Exercise 1.4 Determine the first order partial derivatives of each of the following functions

(a) $f(x, y) = x^3 + x^2y^3 - 2y^2$;

(b) $f(x, y) = \dfrac{x}{(x + y)^2}$;

(c) $f(x, y, z) = xz - 5x^2y^3z^4$;

(d) $f(x, y, z) = ze^{xyz}$;

(e) $f(x, y, z) = \log\left(x + \sqrt{y^2 + z^2}\right)$.

1.2.2 Critical Points

As with functions of one variable an extremum (maximum or minimum) can only occur at a vector where all first order partial derivatives are 0.

We write this as

$$\nabla f(\mathbf{x}) = 0,$$

where

$$\nabla f(\mathbf{x}) = \left(\frac{\partial f}{\partial x_1}(\mathbf{x}), \frac{\partial f}{\partial x_2}(\mathbf{x}), \ldots, \frac{\partial f}{\partial x_n}(\mathbf{x})\right).$$

We call ∇f the gradient of f.

If $n = 3$ we have

$$\nabla f(\mathbf{x}) = \left(\frac{\partial f}{\partial x}(\mathbf{x}), \frac{\partial f}{\partial y}(\mathbf{x}), \frac{\partial f}{\partial z}(\mathbf{x})\right)$$

Let us first consider a function of two variables.

Example 1.5 Let

$$f(x, y) = -2x^2 - y^2 + 8x + 10y - 5xy.$$

We have

$$\frac{\partial f}{\partial x}(x,y) = -4x + 8 - 5y, \qquad \frac{\partial f}{\partial y}(x,y) = -2y + 10 - 5x.$$

If (x, y) is a critical point of f then we have

$$\begin{aligned} 4x + 5y &= 8 \\ 5x + 2y &= 10 \end{aligned}$$

Multiply the first equation by 5 and the second by 4 to get

$$\begin{aligned} 20x + 25y &= 40 \\ 20x + 8y &= 40 \end{aligned}$$

Subtracting we get that $17y = 0$. Hence $y = 0$ and therefore $x = 2$. This means that the point $(2, 0)$ is a critical point of f.

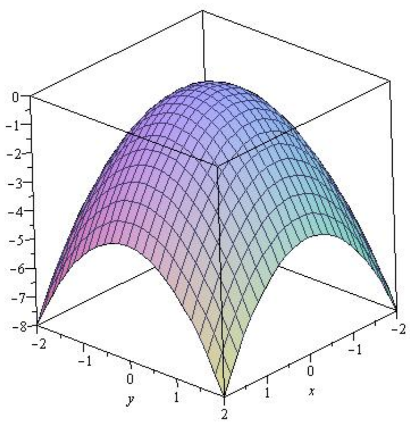
Exercise 1.6 Find the critical points of $f(x, y) = x^2 + y^2 - 2x - 6y + 14$.

Exercise 1.7 Find the critical points of $f(x, y) = x^3 - 12xy + 8y^3$.

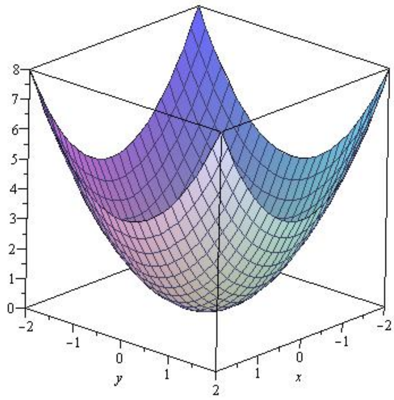
At a critical point, a function can have

- a (local) maximum,
- a (local) minimum,
- a saddle point.

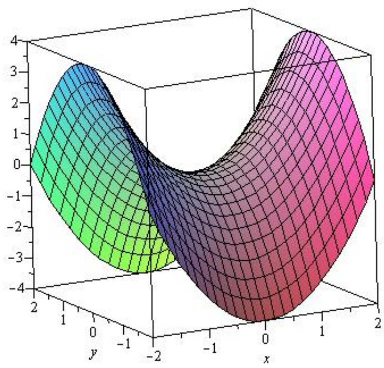
An example of a maximum point:



An example of a minimum point:



An example of a saddle point:



We know that extreme values will occur at critical points. But how do we know the nature of a critical point?

As with functions functions of one variable, to classify critical points of functions of several variables we need higher order partial derivatives.

For functions of two variables, there are four second order partial derivatives:

$$
\begin{aligned}
\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) &= \frac{\partial^2 f}{\partial x^2}, & \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) &= \frac{\partial^2 f}{\partial y \partial x}, \\
\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) &= \frac{\partial^2 f}{\partial x \partial y}, & \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) &= \frac{\partial^2 f}{\partial y^2}.
\end{aligned}
$$

We also use the notation $f_{xy} = \frac{\partial^2 f}{\partial y \partial x}$. Note the order of the indices!

For functions of n variables there are n^2 second order partial derivatives

$$
f_{x_i x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right).
$$

Note that in most of the examples we will consider that $f_{x_i x_j} = f_{x_j x_i}$.

Example 1.8 If $f\colon \mathbb{R}^3 \to \mathbb{R}$ is given by

$$
f(x, y, z) = 5x^2 + 4xy + y^3 - 2y + z^5 + xe^z
$$

then

$$
\begin{array}{ccc}
f_{xx} = 10 & f_{xy} = 4 & f_{xz} = e^z \\
f_{yx} = 4 & f_{yy} = 6y & f_{yz} = 0 \\
f_{xz} = e^z & f_{yz} = 0 & f_{zz} = 20z^3 + xe^z.
\end{array}
$$

Exercise 1.9 Determine the second order partial derivatives of each of the following functions

(a) $f(x, y) = x^3 + x^2y^3 - 2y^2$;

(b) $f(x, y) = \dfrac{x}{(x+y)^2}$;

(c) $f(x, y, z) = xz - 5x^2y^3z^4$;

(d) $f(x, y, z) = ze^{xyz}$;

(e) $f(x, y, z) = \log\left(x + \sqrt{y^2 + z^2}\right)$.

1.3 The Second Derivative Test

We need to introduce some linear algebra.

Otto Hesse 1811-74

1.3.1 The Hessian Matrix

The Hessian matrix H of the function f is the matrix of all second order partial derivatives of f.

$$H(\mathbf{x}) = \begin{pmatrix} f_{x_1x_1}(\mathbf{x}x) & f_{x_1x_2}(\mathbf{x}) & f_{x_1x_3}(\mathbf{x}) & \cdots & f_{x_1x_n}(\mathbf{x}) \\ f_{x_2x_1}(\mathbf{x}) & f_{x_2x_2}(\mathbf{x}) & f_{x_2x_3}(\mathbf{x}) & \cdots & f_{x_2x_n}(\mathbf{x}) \\ f_{x_3x_1}(\mathbf{x}) & f_{x_3x_2}(\mathbf{x}) & f_{x_3x_3}(\mathbf{x}) & \cdots & f_{x_2x_n}(\mathbf{x}) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ f_{x_nx_1}(\mathbf{x}) & f_{x_nx_2}(\mathbf{x}) & f_{x_nx_3}(\mathbf{x}) & \cdots & f_{x_nx_n}(\mathbf{x}) \end{pmatrix},$$

For most functions used in practical applications, we have

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i},$$

hence the Hessian matrix is symmetric.

Example 1.10 The Hessian matrix of $f(x, y, z) = 5x^2 + 4xy + y^3 - 2y + z^5 + xe^z$ is

$$H(\mathbf{x}) = \begin{pmatrix} 10 & 4 & e^z \\ 4 & 6y & 0 \\ e^z & 0 & 20z^3 + xe^z \end{pmatrix}.$$

Exercise 1.11 Determine the Hessian matrix of each of the following functions

(a) $f(x, y) = x^3 + x^2y^3 - 2y^2$;

(b) $f(x,y) = \dfrac{x}{(x+y)^2}$;

(c) $f(x,y,z) = xz - 5x^2y^3z^4$;

(d) $f(x,y,z) = ze^{xyz}$;

(e) $f(x,y,z) = \log\left(x + \sqrt{y^2+z^2}\right)$;

(f) $f(x,y,z) = xe^y + ye^z + ze^x$.

For an $n \times n$ Hessian matrix we identify n submatrices. The first of these, denoted by $H_1(\mathbf{x})$, is the 1×1 submatrix consisting of the element located in row 1 and column 1,

$$H_1(\mathbf{x}) = \begin{pmatrix} f_{x_1x_1}(\mathbf{x}) \end{pmatrix}.$$

The second submatrix is the 2×2 matrix

$$H_2(\mathbf{x}) = \begin{pmatrix} f_{x_1x_1}(\mathbf{x}) & f_{x_1x_2}(\mathbf{x}) \\ f_{x_2x_1}(\mathbf{x}) & f_{x_2x_2}(\mathbf{x}) \end{pmatrix}.$$

The third submatrix is the 3×3 matrix

$$H_3(\mathbf{x}) = \begin{pmatrix} f_{x_1x_1}(\mathbf{x}) & f_{x_1x_2}(\mathbf{x}) & f_{x_1x_3}(\mathbf{x}) \\ f_{x_2x_1}(\mathbf{x}) & f_{x_2x_2}(\mathbf{x}) & f_{x_2x_3}(\mathbf{x}) \\ f_{x_3x_1}(\mathbf{x}) & f_{x_3x_2}(\mathbf{x}) & f_{x_3x_3}(\mathbf{x}) \end{pmatrix},$$

The j^{th} submatrix is the $j \times j$ matrix

$$H_j(\mathbf{x}) = \begin{pmatrix} f_{x_1x_1}(\mathbf{x}) & \cdots & f_{x_1x_j}(\mathbf{x}) \\ \vdots & & \vdots \\ f_{x_jx_1}(\mathbf{x}) & \cdots & f_{x_jx_j}(\mathbf{x}) \end{pmatrix}$$

and $H_n(\mathbf{x}) = H(\mathbf{x})$.

Example 1.12 If $f \colon \mathbb{R}^3 \to \mathbb{R}$ is given by

$$f(x,y,z) = 5x^2 + 4xy + y^3 - 2y + z^5 + xe^z$$

then

$$H(\mathbf{x}) = \begin{pmatrix} 10 & 4 & e^z \\ 4 & 6y & 0 \\ e^z & 0 & 20z^3 + xe^z \end{pmatrix}.$$

So

$$H_1(\mathbf{x}) = 10,$$

$$H_2(\mathbf{x}) = \begin{pmatrix} 10 & 4 \\ 4 & 6y \end{pmatrix}$$

and

$$H_3(\mathbf{x}) = \begin{pmatrix} 10 & 4 & e^z \\ 4 & 6y & 0 \\ e^z & 0 & 20z^3 + xe^z \end{pmatrix}.$$

The determinants of these submatrices are called principal minors and are denoted by $\Delta_k(\mathbf{x})$

$$\Delta_k(\mathbf{x}) = \det(H_k(\mathbf{x})) \qquad k = 1, \ldots, n.$$

Example 1.13 The principal minors of $f(x, y, z) = 5x^2 + 4xy + y^3 - 2y + z^5 + xe^z$ are

$$\Delta_1(\mathbf{x}) = H_1(\mathbf{x}) = 10$$

$$\Delta_2(\mathbf{x}) = \det H_2(\mathbf{x}) = 60y - 16$$

$$\Delta_3(\mathbf{x}) = \det H_3(\mathbf{x}) = e^z(-6ye^z) + (20z^3 + xe^z)(60y - 16)$$

Exercise 1.14 Determine the principal minors of each of the following functions

(a) $f(x, y, z) = xz - 5x^2y^3z^4$;

(b) $f(x, y, z) = ze^{xyz}$;

(c) $f(x, y, z) = \log(x + \sqrt{y^2 + z^2})$.

1.3.2 Second Derivative Test

Theorem 1.15 (Second Derivative Test) Let $\mathbf{p}$ be a critical point of a function $f\colon \mathbb{R}^n \to \mathbb{R}$.

- If $(-1)^k \Delta_k(\mathbf{p}) > 0$ for all $k = 1, \ldots, n$, then f has a local maximum at $\mathbf{p}$.
- If $\Delta_k(\mathbf{p}) > 0$ for all $k = 1, \ldots, n$, then f has a local minimum at $\mathbf{p}$.
- If $\Delta_n(\mathbf{p}) = \det(H(\mathbf{p})) \neq 0$, but none of the previous conditions holds, then f has a saddle point at $\mathbf{p}$.
- If $\Delta_n(\mathbf{p}) = \det(H(\mathbf{p})) = 0$, the test is inconclusive.

Exercise 1.16 Find and classify all critical points of the function

$$f(x, y, z) = 2x^2 + xy + 4y^2 + xz + z^2 + 2.$$

We have

$$f_x(x, y, z) = 4x + y + z, \qquad f_y(x, y, z) = x + 8y, \qquad f_z(x, y, z) = x + 2z.$$

Let p be a critical point of f. Then we have

$$\begin{aligned} f_x(x, y, z) = 0 &\Rightarrow 4x + y + z = 0 \\ f_y(x, y, z) = 0 &\Rightarrow x + 8y = 0 \\ f_z(x, y, z) = 0 &\Rightarrow x + 2z = 0 \end{aligned}$$

So $4x - \frac{1}{8}x - \frac{1}{2}x = 0$ which gives $3\frac{3}{8}x = 0$ and thus $x = 0$. This in turn gives $y = 0$ and $z = 0$. Therefore, the only critical point of f is $(0, 0, 0)$.

The Hessian matrix of f is given by

$$H_f(\mathbf{x}) = \begin{pmatrix} 4 & 1 & 1 \\ 1 & 8 & 0 \\ 1 & 0 & 2 \end{pmatrix}.$$

We have that the principal minors of f are

$$\Delta_1(\mathbf{x}) = 4, \qquad \Delta_2(\mathbf{x}) = 31, \qquad \Delta_3(\mathbf{x}) = 1(-8) + 2(31) = 54.$$

As $\Delta_1(\mathbf{x}) > 0$, $\Delta_2(\mathbf{x}) > 0$ and $\Delta_3(\mathbf{x}) > 0$, f has a local minimum at $(0, 0, 0)$.

The following illustrates out procedure in finding and classifying critical points of functions of several variables.

Function f of n variables

$\downarrow$

Determine first order derivatives of f

$\swarrow$ $\searrow$

Set $\nabla f = 0$ and solve | Find second order partial derivatives f

$\searrow$ $\swarrow$

Determine Hessian Matrix H_f

$\downarrow$

Evaluate H_f at critical points

$\downarrow$

Evaluate Principal Minors at critical points

$\downarrow$

Use Second Derivative Test to classify critical points

Example 1.17 Find and classify the critical points of

$$f(x,y,z) = x^2 + y^3 + z^2 - 2xy - 2yz + 1.$$

Finding Critical Points:

We have

$$f_x(x,y,z) = 2x - 2y, \qquad f_y(x,y,z) = 3y^2 - 2x - 2z,$$
$$f_z(x,y,z) = 2z - 2y.$$

Setting $f_x(x,y,z) = 0$ we get $x = y$, $f_z(x,y,z) = 0$ we get $z = y$. So $x = y = z$. Setting $f_z(x,y,z) = 0$ we get $3y^2 = 4y$ so $y = 0$ or $y = \frac{4}{3}$. Therefore we have two critical points, $(0,0,0)$ and $(\frac{4}{3}, \frac{4}{3}, \frac{4}{3})$.

Classifying the Critical Points:

We have

$$\begin{aligned} f_{xx}(x,y,z) &= 2, & f_{xy}(x,y,z) &= -2, & f_{xz}(x,y,z) &= 0, \\ f_{yx}(x,y,z) &= -2, & f_{yy}(x,y,z) &= 6y, & f_{yz}(x,y,z) &= -2, \\ f_{zx}(x,y,z) &= 0, & f_{zy}(x,y,z) &= -2, & f_{zz}(x,y,z) &= 2. \end{aligned}$$

This gives a Hessian matrix of

$$H_f(x,y,z) = \begin{pmatrix} 2 & -2 & 0 \\ -2 & 6y & -2 \\ 0 & -2 & 2 \end{pmatrix}.$$

Therefore

$$H_f(0,0,0) = \begin{pmatrix} 2 & -2 & 0 \\ -2 & 0 & -2 \\ 0 & -2 & 2 \end{pmatrix}$$

giving $\Delta_1(0,0,0) = 2$, $\Delta_2(0,0,0) = -4$, $\Delta_3(0,0,0) = -16$. So by the Second Derivative Test f has a saddle point at $(0,0,0)$.

Also,

$$H_f\left(\frac{4}{3},\frac{4}{3},\frac{4}{3}\right) = \begin{pmatrix} 2 & -2 & 0 \\ -2 & 8 & -2 \\ 0 & -2 & 2 \end{pmatrix}$$

giving $\Delta_1(\frac{4}{3},\frac{4}{3},\frac{4}{3}) = 2$, $\Delta_2(\frac{4}{3},\frac{4}{3},\frac{4}{3}) = 12$, $\Delta_3(\frac{4}{3},\frac{4}{3},\frac{4}{3}) = 16$. So by the Second Derivative Test f has a local minimum at $(\frac{4}{3},\frac{4}{3},\frac{4}{3})$.

Example 1.18 Find and classify all critical points of the function

$$f(x,y,z) = x^2 + x^2y + y^2z + z^2 - 4z.$$

We have

$$f_x(x,y,z) = 2x + 2xy, \qquad f_y(x,y,z) = x^2 + 2yz, \qquad f_z(x,y,z) = y^2 + 2z - 4.$$

Let p be a critical point of f. Then we have

$$\begin{aligned} f_x(x,y,z) = 0 &\Rightarrow 2x + 2xy = 0 \\ f_y(x,y,z) = 0 &\Rightarrow x^2 + 2yz = 0 \\ f_z(x,y,z) = 0 &\Rightarrow y^2 + 2z - 4 = 0. \end{aligned}$$

Multiply the first equation by z, the second equation by x and the third by xy to get

$$\begin{aligned} 2xz + 2xyz &= 0 \\ x^3 + 2xyz &= 0 \\ xy^3 + 2xyz - 4xy &= 0. \end{aligned}$$

Now subtract the second equation from the first, the third from the second and the third from the first to get

$$\begin{aligned} 2xz - x^3 &= 0 \\ x^3 - xy^3 + 4xy &= 0 \\ 2xz - xy^3 + 4xy &= 0. \end{aligned}$$

Therefore we have $x(x^2 - 2z) = 0$ so either $x = 0$ or $z = \frac{x^2}{2}$.

If $x = 0$ then $2yz = 0$ and hence $y = 0$ or $z = 0$. If $y = 0$ then $y^2 + 2z - 4$ gives $z = 2$. If $z = 0$ then $y^2 + 2z - 4$ gives $y = \pm 2$. This gives us critical points of $(0, 2, 0)$, $(0, -2, 0)$ and $(0, 0, 2)$.

If $z = \frac{x^2}{2}$ and $x \neq 0$ then $2x(1 + y) = 0$ gives $y = -1$. This now gives $(-1)^2 + 2z - 4 = 0$ giving $z = \frac{3}{2}$. The equation $x^2 = 2z$ now gives $x = \pm\sqrt{3}$ giving two new critical points of $(\sqrt{3}. - 1, \frac{3}{2})$ and $(-\sqrt{3}. - 1, \frac{3}{2})$.

The second order partial derivatives of f are

$$\begin{aligned} f_{xx} &= 2 + 2y & f_{xy} &= 2x & f_{xz} &= 0 \\ f_{yx} &= 2x & f_{yy} &= 6z & f_{yz} &= 2y \\ f_{zx} &= 0 & f_{zy} &= 2y & f_{zz} &= 2. \end{aligned}$$

and therefore the Hessian of f at (x, y, z) is given by

$$H_f(\mathbf{x}) = \begin{pmatrix} 2 + 2y & 2x & 0 \\ 2x & 6z & 2y \\ 0 & 2y & 2 \end{pmatrix}.$$

We now check each of our critical points.

Let us start with $(0, 2, 0)$. We have

$$H_f((0, 2, 0)) = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 0 & 4 \\ 0 & 4 & 2 \end{pmatrix}.$$

Then $\Delta_1(0, 2, 0) = 6$, $\Delta_2(0, 2, 0) = 0$ and $\Delta_3(0, 2, 0) = -96$. Therefore f has a saddle point at $(0, 2, 0)$.

Next, we consider the point $(0, -2, 0)$. We have

$$H_f((0, -2, 0)) = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 0 & -4 \\ 0 & -4 & 2 \end{pmatrix}.$$

Then $\Delta_1(0, -2, 0) = -2$, $\Delta_2(0, -2, 0) = 0$ and $\Delta_3(0, -2, 0) = 32$. Again f has a saddle point at $(0, -2, 0)$.

Next, we consider the critical point $(0, 0, 2)$. We have

$$H_f((0, 0, 2)) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 12 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Then $\Delta_1(0,0,2) = 2$, $\Delta_2(0,0,2) = 24$ and $\Delta_3(0,0,2) = 48$. Therefore f has a local minimum at $(0,0,2)$.

Next, we consider the critical point $(\sqrt{3}, -1, \frac{2}{2})$. We have

$$H_f((\sqrt{3}, -1, \tfrac{3}{2})) = \begin{pmatrix} 0 & 2\sqrt{3} & 0 \\ 2\sqrt{3} & 9 & -2 \\ 0 & -2 & 2 \end{pmatrix}.$$

Then $\Delta_1(\sqrt{3}, -1, \frac{3}{2}) = 0$, $\Delta_2(\sqrt{3}, -1, \frac{3}{2}) = -12$ and $\Delta_3(\sqrt{3}, -1, \frac{3}{2}) = 24$. Therefore f has a saddle point at $(\sqrt{3}, -1, \frac{3}{2})$.

Finally, let us consider the critical point $(-\sqrt{3}, -1, \frac{2}{2})$. We have

$$H_f(-\sqrt{3}, -1, \tfrac{3}{2})) = \begin{pmatrix} 0 & -2\sqrt{3} & 0 \\ -2\sqrt{3} & 9 & -2 \\ 0 & -2 & 2 \end{pmatrix}.$$

Then $\Delta_1(-\sqrt{3}, -1, \frac{3}{2}) = 0$, $\Delta_2(-\sqrt{3}, -1, \frac{3}{2}) = -12$ and $\Delta_3(-\sqrt{3}, -1, \frac{3}{2}) = 24$. And again f has a saddle point at $(-\sqrt{3}, -1, \frac{2}{2})$.

Example 1.19 For $a \in \mathbb{R}$ and consider the function

$$f(x,y,z) = x^2 + ay^2 + z^2 - 4xy.$$

Determine how the critical points of f and their nature depends on a.

Finding Critical Points:

We have

$$f_x(x,y,z) = 2x - 4y, \qquad f_y(x,y,z) = 2ay - 4x,$$
$$f_z(x,y,z) = 2z.$$

We see immediately from $f_z(x,y,z) = 0$ that $z = 0$.

Setting $f_x(x,y,z) = 0$ we get $x = 2y$. Setting $f_y(x,y,z) = 0$ we get $ay = 2x$. Therefore we have that $ay = 4y$ and hence either have $y = 0$ or $a = 4$. If $y = 0$ then $x = 0$ and we get the critical point $(0,0,0)$. If $a = 4$ we that any point of the form $(2y, y, 0)$ is a critical point of f.

Classifying the Critical Points:

We have

$$f_{xx}(x,y,z) = 2, \qquad f_{xy}(x,y,z) = -4, \qquad f_{xz}(x,y,z) = 0,$$
$$f_{yx}(x,y,z) = -4, \qquad f_{yy}(x,y,z) = 2a, \qquad f_{yz}(x,y,z) = 0,$$

$$f_{zx}(x,y,z) = 0, \qquad f_{zy}(x,y,z) = 0, \qquad f_{zz}(x,y,z) = 2.$$

This gives a Hessian matrix of

$$H_f(x,y,z) = \begin{pmatrix} 2 & -4 & 0 \\ -4 & 2a & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Therefore

$$H_f(0,0,0) = \begin{pmatrix} 2 & -4 & 0 \\ -4 & 2a & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

giving $\Delta_1(0,0,0) = 2$, $\Delta_2(0,0,0) = 4a - 16$, $\Delta_3(0,0,0) = 2(4a - 16)$. So by the Second Derivative Test f has a local minimum at $(0,0,0)$ if $a > 4$ and a saddle point $(0,0,0)$ if $a < 4$.

If $a = 4$ the Second Derivative Test is inconclusive.

When $a = 4$ we get critical points of the form $(2y, y, 0)$ for $y \in \mathbb{R}$. In this case we see that

$$H_f(2y,y,0) = \begin{pmatrix} 2 & -4 & 0 \\ -4 & 8 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

giving $\Delta_1(2y,y,0) = 2$, $\Delta_2(2y,y,0) = 0$, $\Delta_3(2y,y,0) = 0$ which means that the Second Derivative Test is inconclusive.

When $a = 4$ we see that every points on the line $\{(2y,y,0) : y \in \mathbb{R}\}$ are all critical points of f. However, as we observe that $f(2y,y,0) = 0$ points of the form $(2y,y,0)$ are neither local maximum, local minimum of saddle points.

Exercise 1.20 Find and classify the critical points of the following functions

(a) $f(x,y,z) = x^2 + y^2 + 7z^2 - xy - 3yz$;

(b) $f(x,y,z) = x^4 + (x+y)^2 + (x+z)^2$;

(c) $f(x,y,z) = x^4 + x^2y + y^2 + z^2 + xy + 1$;

(d) $f(x,y,z) = x^3 - 2x^2 + y^2 + z^2 - 2xy + xz - yz + 3z$;

(e) $f(x,y,z,w) = -x^2 - y^2 - z^2 - w^2 + xy + wz$.

Keywords: Functions of Several Variables; Partial Derivative; Critical Point; Hessian Matrix; Second Derivatives Test.

Chapter 2

Lagrange Multipliers

2.1 Lagrange Multipliers with one constraint

[1]

Let us recall the method of Lagrange multiplies for functions of two variables and one constraint

> Problem 2.1 Find the dimensions of the rectangle with maximum area, given that the perimeter is 10 m.

We can translate this into a mathematical problem:

> Problem 2.2 For what (x, y) is the function $f(x, y) = xy$ maximal, given the constraint $g(x, y) = 2x + 2y = 10$.

The Lagrangian method tells us, that if we want to find the maximum (minimum) of $f(x, y)$ subject to $g(x, y) = c$,

We construct the Lagrangian Function

$$F(x, y, \lambda) = f(x, y) + \lambda[c - g(x, y)],$$

determine its critical points by setting

$$F_x(x, y, \lambda) = 0,$$

$$F_y(x, y, \lambda) = 0,$$

$$F_\lambda(x, y, \lambda) = 0$$

and solving.

The variable λ is known as the Lagrange Multiplier.

For the above problem we let $g(x,y) = 2x + 2y$ and set

$$\begin{aligned} F(x,y,\lambda) =& f(x,y) + \lambda[10 - g(x,y)] \\ =& xy + \lambda[10 - 2x - 2y] \end{aligned}$$

Then

$$\begin{aligned} F_x(x,y,\lambda) =& y + \lambda[-2], \\ F_y(x,y,\lambda) =& x + \lambda[-2], \\ F_\lambda(x,y,\lambda) =& 10 - 2x - 2 - y. \end{aligned}$$

Setting each of these equal to 0 we see that

$$\lambda = \frac{x}{2} = \frac{y}{2}$$

giving $x = y$. Putting this in the equation $2x + 2y = 10$ we get that $x = y = 2\frac{1}{2}$ and $\lambda = 1\frac{1}{4}$.

There are two 'challenges' with using the method of Lagrange Multipliers. The first is that there is no systematic way to solve the set of equations we obtain when we compute the partial derivatives. The second is that, unlike the Second Derivative Test, the method of Lagrange Multipliers does not tell us if our solution is a maximum, minimum or saddle point. We distinguish between them by looking at the relative values the objective function takes at the critical points.

Example 2.3 Find the maximum (minimum) of $f(x,y) = x^2+y^2$ subject to $x^2 + xy + y^2 = 3$.

We let $g(x,y) = x^2 + xy + y^2$ and set

$$\begin{aligned} F(x,y,\lambda) =& f(x,y) + \lambda[3 - g(x,y)] \\ =& x^2 + y^2 + \lambda[3 - x^2 - xy - y^2] \end{aligned}$$

Then

$$\begin{aligned} F_x(x,y,\lambda) =& 2x + \lambda[-2x - y], \\ F_y(x,y,\lambda) =& 2y + \lambda[-x - 2y], \\ F_\lambda(x,y,\lambda) =& 3 - x^2 - xy - y^2. \end{aligned}$$

Setting each of these equal to 0 we obtain the equations

$$\begin{aligned} 2x + \lambda[-2x - y] =& 0, \\ 2y + \lambda[-x - 2y] =& 0, \\ 3 - x^2 - xy - y^2 =& 0. \end{aligned}$$

Note that $2x + y \neq 0$ as $2x + y = 0$ implies $x = 0$ and therefore $y = 0$. This is not possible as $x^2 + xy + y^2 = 3$.

Similary, $x + 2y \neq 0$.

From the first and second we get that

$$\lambda = \frac{2x}{2x + y} = \frac{2y}{2y + x}.$$

So

$$\begin{aligned} 2x(2y + x) =& 2y(2y + x) \\ &\Rightarrow 2x^2 = 2y^2 \\ &\Rightarrow x = \pm y \end{aligned}$$

If $x = y$ then $3x^2 = 3$ giving $x = \pm 1$ and $y = \pm 1$.

If $x = -y$ then $x^2 = 3$ giving $x = \pm\sqrt{3}$ and $y = \mp\sqrt{3}$.

This gives us four possible extreme points $(1, 1)$, $(-1, -1)$, $(\sqrt{3}, -\sqrt{3})$ and $(-\sqrt{3}, \sqrt{3})$. We have $f(1, 1) = f(-1, -1) = 2$ while $f(\sqrt{3}, -\sqrt{3}) = f(-\sqrt{3}, \sqrt{3}) = 6$.

Hence f has a maximum at 6 which occurs at $(\sqrt{3}, -\sqrt{3})$ and $(-\sqrt{3}, \sqrt{3})$ and a minimum of 1 which occurs at $(1, 1)$ and $(-1, -1)$.

Example 2.4 Consider the Cobb-Douglas production function

$$q(k, l) = 25k^{1/3}l^{1/6}$$

where q, k and l denote quantity, labour and capital respectively in a certain industrial process. Maximise production subject to $10k+5l = 30$.

Let us set $f(k, l) = 25k^{1/3}l^{1/6}$, $g(k, l) = 10k + 5l$ and $c = 30$. Then our Lagrangian function is

$$F(k, l, \lambda) = 25k^{1/3}l^{1/6} + \lambda[30 - 10k - 5l].$$

We have

$$\begin{aligned} F_k(k, l, \lambda) =& \frac{25}{3}k^{-2/3}l^{1/6} - 10\lambda, \\ F_l(k, l, \lambda) =& \frac{25}{6}k^{1/3}l^{-5/6} - 5\lambda, \\ F_\lambda(k, l, \lambda) =& 30 - 10k - 5l. \end{aligned}$$

Setting each of these equal to 0 we get

$$\begin{aligned}\frac{25}{3}k^{-2/3}l^{1/6} - 10\lambda =&0.\\ \frac{25}{6}k^{1/3}l^{-5/6} - 5\lambda =&0.\\ 30 - 10k + 5\lambda =&0.\end{aligned}$$

This gives that

$$\lambda = \frac{5}{6}k^{-2/3}l^{1/6} = \frac{5}{6}\frac{\sqrt[6]{l}}{\sqrt[3]{k^2}}$$

and

$$\lambda = \frac{5}{6}k^{1/3}l^{-5/6} = \frac{5}{6}\frac{\sqrt[3]{k}}{\sqrt[6]{l^5}}.$$

So

$$\lambda = \frac{5}{6}\frac{\sqrt[3]{k}}{\sqrt[6]{l^5}} = \frac{5}{6}\frac{\sqrt[6]{l}}{\sqrt[3]{k^2}}.$$

This gives $\sqrt[3]{k^3} = \sqrt[6]{l^6}$ or that $k = l$.

Putting this into $10k + 5l = 30$ we get $15k = 30$. So $k = l = 2$ and $\lambda = \frac{5}{6}\frac{\sqrt[3]{2}}{\sqrt[6]{2^5}} = \frac{5}{6}\frac{1}{\sqrt{2}}$ and $q = \sqrt{2}(25)$.

Exercise 2.5 Use the method of Lagrange Multipliers to find

(a) the maximum and minimum values of $f(x, y) = x^2 + y^2$ subject to $xy = 1$,

(b) the maximum and minimum values of $f(x, y) = 3x + y$ subject to $x^2 + y^2 = 10$,

(c) the maximum and minimum values of $f(x, y) = y^2 - x^2$ subject to $\frac{1}{4}x^2 + y^2 = 1$,

(d) the maximum and minimum values of $f(x, y) = e^{xy}$ subject to $x^2 + y^2 = 16$.

2.2 Lagrangian method with n variables and one constraint

Suppose we have a function $f\colon \mathbb{R}^n \to \mathbb{R}$ and we want to find the maximum (minimum) of $f(x_1, \ldots, x_n)$ subject to $g(x_1, \ldots, x_n) = c$. We consider the Lagrangian function

$$F(x_1, \ldots, x_n, \lambda) = f(x_1, \ldots, x_n) + \lambda[c - g(x_1, \ldots, x_n]).$$

Joseph-Louis Lagrange 1736-1813

We then solve the $n+1$ equations

$$\begin{aligned}
F_{x_1}(x_1,\ldots,x_n,\lambda) &= 0\\
F_{x_2}(x_1,\ldots,x_n,\lambda) &= 0\\
\vdots \qquad & \quad \vdots\\
F_{x_n}(x_1,\ldots,x_n,\lambda) &= 0\\
F_{\lambda}(x_1,\ldots,x_n,\lambda) &= 0
\end{aligned}$$

Example 2.6 Find the maximum of xyz subject to $x+2y+4z=12$.

Let us set $g(x,y,z)=x+2y+4z$ and $c=12$. Then our Lagrangian function is

$$F(x,y,z,\lambda)=xyz+\lambda[12-x-2y-4z].$$

We have

$$\begin{aligned}
F_x(x,y,z,\lambda) &= yz-\lambda,\\
F_y(x,y,z,\lambda) &= xz-2\lambda,\\
F_z(x,y,z,\lambda) &= xy-4\lambda,\\
F_\lambda(x,y,z,\lambda) &= 12-x-2y-4z.
\end{aligned}$$

Setting each of these equal to 0 we get

$$\begin{aligned} yz - \lambda =& 0, \\ xz - 2\lambda =& 0, \\ xy - 4\lambda =& 0 \\ 12 - x - 2y - 4z =& 0. \end{aligned}$$

Comparing values for λ we get

$$\lambda = yz = \frac{xz}{2} = \frac{xy}{4}.$$

For a maximum we must have $x \neq 0$, $y \neq 0$ and $z \neq 0$. So we can divide by y and z to get that

$$z = \frac{x}{4}, \qquad y = \frac{x}{2}.$$

As $x + 2y + 4z = 12$ we get that $3x = 12$ and therefore $x = 4$, $y = 2$, $z = 1$ and $\lambda = 1$. Therefore a maximum of 8 occur when $x = 4$, $y = 2$ and $z = 1$.

Example 2.7 Suppose

$$U(x, y, z) = xyz$$

is the utility function of a person consuming x, y and z units of three commodities X, Y and Z.

Suppose that X costs €1 per unit, Y costs €4 per unit and Z costs €2 unit.

(a) If a person has a budget of €48, how much of each unit should he or she buy in order to maximise utility?

(b) What is the maximum utility?

In mathematical terms the problem is to find the maximum of $U(x, y, z) = xyz$ subject to to $x + 4y + 2z = 48$.

Let us set $g(x, y, z) = x + 4y + 2z$ and $c = 48$. Then our Lagrangian function is

$$F(x, y, z, \lambda) = xyz + \lambda[48 - x - 4y - 2z].$$

We have

$$\begin{aligned} F_x(x, y, z, \lambda) =& yz - \lambda, \\ F_y(x, y, z, \lambda) =& xz - 4\lambda, \\ F_z(x, y, z, \lambda) =& xy - 2\lambda, \\ F_\lambda(x, y, z, \lambda) =& 48 - x - 4y - 2z. \end{aligned}$$

Setting each of these equal to 0 we get

$$\begin{aligned}
yz - \lambda =& 0, \\
xz - 4\lambda =& 0, \\
y - 2\lambda =& 0 \\
48 - x - 4y - 2z =& 0.
\end{aligned}$$

Comparing values for λ we get

$$\lambda = yz = \frac{xz}{4} = \frac{xy}{2}.$$

For a maximum we must have $x \neq 0$, $y \neq 0$ and $z \neq 0$. So we can divide by y and z to get that

$$z = \frac{x}{2}, \qquad y = \frac{x}{4}.$$

As $x + 4y + 2z = 48$ we get that $3x = 48$ and therefore $x = 16$, $y = 4$, $z = 8$ and $\lambda = 32$. Therefore a maximum of 512 occur when $x = 16$, $y = 4$ and $z = 8$.

Exercise 2.8 Use the method of Lagrange Multipliers to find

(a) the maximum and minimum values of $f(x, y, z) = x^2 + y^2 + z^2$ subject to $x + y + z = 12$,

(b) the maximum and minimum values of $f(x, y, z) = 2x + 2y + z$ subject to $x^2 + y^2 + z^2 = 9$,

(c) the maximum and minimum values of $f(x, y, z) = x^2y^2z^2$ subject to $x^2 + y^2 + z^2 = 1$,

(d) the maximum and minimum values of $f(x, y, z) = x^4 + y^4 + z^4$ subject to $x^2 + y^2 + z^2 = 1$,

(e) the maximum and minimum values of $f(x, y, z, t) = x + y + z + t$ subject to $x^2 + y^2 + z^2 + t^2 = 1$.

2.3 The Lagrangian Method: n variables, m constraints

Consider the following problem: Find the maximum(minimum) of $f(x_1, \ldots, x_n)$ subject to $g_1(x_1, \ldots, x_n) = c_1, g_2(x_1, \ldots, x_n) = c_2, \ldots, g_m(x_1, \ldots, x_n) = c_m$.

We consider the Lagrangian function

$$\begin{aligned}
F(x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_m) =& f(x_1, \ldots, x_n) + \lambda_1[c_1 - g_1(x_1, \ldots, x_n)] \\
&+ \lambda_2[c_2 - g_2(x_1, \ldots, x_n)] + \ldots \\
&+ \lambda_m[c_m - g_m(x_1, \ldots, x_n)].
\end{aligned}$$

To find the maximum(minimum) of $f(x_1,\dots,x_n)$ subject to $g_1(x_1,\dots,x_n) = c_1, g_2(x_1,\dots,x_n) = c_2,\dots,g_m(x_1,\dots,x_n) = c_m$ we solve the $n+m$ equations

$$\begin{aligned}
F_{x_1}(x_1,\dots,x_n,\lambda_1,\dots,\lambda_m) =&0,\\
F_{x_2}(x_1,\dots,x_n,\lambda_1,\dots,\lambda_m) =&0,\\
\vdots \qquad\qquad & \vdots\\
F_{x_n}(x_1,\dots,x_n,\lambda_1,\dots,\lambda_m) =&0,\\
F_{\lambda_1}(x_1,\dots,x_n,\lambda_1,\dots,\lambda_m) =&0,\\
F_{\lambda_2}(x_1,\dots,x_n,\lambda_1,\dots,\lambda_m) =&0,\\
\vdots \qquad\qquad & \vdots\\
F_{\lambda_m}(x_1,\dots,x_n,\lambda_1,\dots,\lambda_m) =&0.
\end{aligned}$$

Example 2.9 Find the maximum (minimum) of $f(x,y,z) = x^2+y^2+z^2$ subject to

$$\begin{aligned}
x+2y+z &= 1,\\
2x-y-3z &= 4.
\end{aligned}$$

Our Lagrangian function is

$$F(x,y,z,\lambda_1,\lambda_2) = x^2+y^2+z^2+\lambda_1[1-x-2y-z]+\lambda_2[4-2x+y+3z].$$

Then we have

$$\begin{aligned}
F_x(x,y,z,\lambda_1,\lambda_2) =&2x-\lambda_1-2\lambda_2,\\
F_y(x,y,z,\lambda_1,\lambda_2) =&2y-2\lambda_1+\lambda_2,\\
F_z(x,y,z,\lambda_1,\lambda_2) =&2z-\lambda_1+3\lambda_2,\\
F_{\lambda_1}(x,y,z,\lambda_1,\lambda_2) =&1-x-2y-z,\\
F_{\lambda_2}(x,y,z,\lambda_1,\lambda_2) =&4-2x+y+3z.
\end{aligned}$$

Setting each of these equal to 0 we get

$$\begin{aligned}
2x &= \lambda_1+2\lambda_2,\\
2y &= 2\lambda_1-\lambda_2,\\
2z &= \lambda_1-3\lambda_2,\\
x+2y+z &= 1,\\
2x-y-3z &= 4.
\end{aligned}$$

So, replacing x, y and z with λ_1 and λ_2 we get

$$\begin{aligned}
(\lambda_1+2\lambda_2)+2(2\lambda_1-\lambda_2)+(\lambda_1-3\lambda_2) =&2,\\
2(\lambda_1+2\lambda_2)-(2\lambda_1-\lambda_2)-3(\lambda_1-3\lambda_2) =&8.
\end{aligned}$$

or

$$
\begin{aligned}
6\lambda_1 - 3\lambda_2 =& 2,\\
-3\lambda_1 + 14\lambda_2 =& 8.
\end{aligned}
$$

This gives that $25\lambda_2 = 18$ and hence we have that $\lambda_2 = 18/25$ and $\lambda_1 = 52/75$. Using the above equations we get that f has a minimum value of $134/15$ at

$$x = 16/15, \qquad y = 1/3, \quad \text{and} \quad z = -/15.$$

Example 2.10 Find the maximum (minimum) of $f(x, y, z) = x$ subject to .

$$
\begin{aligned}
z &= x + y,\\
x^2 + 2y^2 + 2z^2 &= 8.
\end{aligned}
$$

We let $g_1(x, y, z) = x + y - z$ and $g_2(x, y, z) = x^2 + 2y^2 + 2z^2$. Then we have to find the extreme values of $f(x, y, z) = x$ subject to $g_1(x, y, z) = 0$ and $g_2(x, y, z) = 8$.

Our Lagrangian function is

$$F(x, y, z, \lambda_1, \lambda_2) = x + \lambda_1[z - x - y] + \lambda_2[8 - x^2 - 2y^2 - 2z^2].$$

Then we have

$$
\begin{aligned}
F_x(x, y, z, \lambda_1, \lambda_2) =& 1 - \lambda_1 - 2\lambda_2 x,\\
F_y(x, y, z, \lambda_1, \lambda_2) =& -\lambda_1 - 4\lambda_2 y,\\
F_z(x, y, z, \lambda_1, \lambda_2) =& \lambda_1 - 4\lambda_2 z,\\
F_{\lambda_1}(x, y, z, \lambda_1, \lambda_2) =& z - x - y,\\
F_{\lambda_2}(x, y, z, \lambda_1, \lambda_2) =& 8 - x^2 - 2y^2 - 2z^2.
\end{aligned}
$$

Setting each of these equal to 0 we get

$$
\begin{aligned}
\lambda_1 + 2\lambda_2 x &= 1,\\
\lambda_1 + 4\lambda_2 y &= 0,\\
\lambda_1 - 4\lambda_2 z &= 0,\\
z &= x + y,\\
x^2 + 2y^2 + 2z^2 &= 8.
\end{aligned}
$$

If $\lambda_2 = 0$ then we would get that $\lambda_1 = 1$ and $\lambda_1 = 0$ which is clearly impossible. Therefore $\lambda_2 \neq 0$ and as

$$\lambda_1 = -4\lambda_2 y = 4\lambda_2 z$$

we get that $y = -z$. Hence $x = z - y = 2z$ and $x^2 + 2y^2 + 2z^2 = 8$ gives $8z^2 = 8$ and therefore $z = \pm 1$.

When $z = 1$ we get $x = 2$, $y = -1$ and $f(2, -1, 1) = 2$.

When $z = -1$ we get $x = -2$, $y = 1$ and $f(-2, 1, -1) = -2$.

Therefore f has a minimum value of -2 at $(-2, 1, -1)$ and a maximum value of 2 at $(2, -1, 1)$.

Exercise 2.11 Use the method of Lagrange Multipliers to find

(a) the maximum and minimum values of $f(x, y, z) = x + 2y$ subject to $x + y + z = 1$ and $y^2 + z^2 = 4$,

(b) the maximum and minimum values of $f(x, y, z) = 3x - y - 3z$ subject to $x + y - z = 0$ and $x^2 + 2z^2 = 1$,

(c) the maximum and minimum values of $f(x, y, z) = yz + xy$ subject to $xy = 1$ and $y^2 + z^2 = 1$,

(d) the maximum and minimum values of $f(x, y, z) = x^2 + y^2 + z^2$ subject to $x - y = 1$ and $y^2 - z^2 = 4$,

(e) the maximum and minimum values of $f(x, y, z) = x + 2y + 3z$ subject to $x - y + z = 1$ and $x^2 + y^2 = 1$.