Vv255 Lecture 13

Dr Jing Liu

UM-SJTU Joint Institute

June 21, 2017

Definition

Let c be a point in the domain $\mathcal{D}$ of a function $y = f(x)$. Then $f(c)$ is a

- *global/absolute* maximum of f for a set $\mathcal{I} \subset \mathcal{D}$ if

$$f(c) \geq f(x) \qquad \text{for all } x \in \mathcal{I}.$$

- *global/absolute* minimum of f for a set $\mathcal{I} \subset \mathcal{D}$ if

$$f(c) \leq f(x) \qquad \text{for all } x \in \mathcal{I}.$$

- local/relative maximum of f if there is a neighborhood $\mathcal{U} \subset \mathcal{D}$ of c such that

$$f(c) \geq f(x) \qquad \text{for all } x \in \mathcal{U}.$$

- local/relative minimum of f if there is a neighborhood $\mathcal{U} \subset \mathcal{D}$ of c such that

$$f(c) \leq f(x) \qquad \text{for all } x \in \mathcal{U}.$$

- We say f has an extremum at c if f has a maximum or a minimum at c.

Definition

Let (a, b) be a point in the domain $\mathcal{D}$ of a function $z = f(x, y)$. Then $f(a, b)$ is a

- *global/absolute* maximum of $f(x, y)$ for a set $\mathcal{S} \subset \mathcal{D}$ if

$$f(a,b) \geq f(x,y) \qquad \text{for all } (x,y) \in \mathcal{S}.$$

- *global/absolute* minimum of $f(x, y)$ for a set $\mathcal{S} \subset \mathcal{D}$ if

$$f(a,b) \leq f(x,y) \qquad \text{for all } (x,y) \in \mathcal{S}.$$

- local/relative maximum of f if there is a neighbourhood $\mathcal{U} \subset \mathcal{D}$ of (a, b)

$$f(a,b) \geq f(x,y) \qquad \text{for all } (x,y) \in \mathcal{U}.$$

- local/relative minimum of f if there is a neighbourhood $\mathcal{U} \subset \mathcal{D}$ of (a, b)

$$f(a,b) \leq f(x,y) \qquad \text{for all } (x,y) \in \mathcal{U}.$$

- We say f has an extremum at P if f has a maximum or a minimum at P.

- To find the local extreme values of a function of a single variable, we look for
- critical points $\begin{cases} 1. & f' \text{ does not exist.} \\ 2. & f' = 0. \end{cases}$
- At such points, we then try determining their nature:

  local maxima, local minima, or points of inflection.

- The first derivative test for a function of a single variable.

$$\begin{aligned} \text{positive} \quad f'(c) \quad \text{negative} &\implies \text{local maximum,} \\ \text{negative} \quad f'(c) \quad \text{positive} &\implies \text{local minimum,} \\ \text{No sign change} &\implies \text{not a local maximum or local minimum} \end{aligned}$$

- The second derivative test for at point c such $f'(c) = 0$.

$$\begin{aligned} f''(c) \quad \text{is positive} &\implies \text{local minimum at } c \\ f''(c) \quad \text{is negative} &\implies \text{local maximum at } c \\ f''(c) = 0 &\implies \text{inconclusive} \end{aligned}$$

- For functions of several variables, the approach is similar. We look for
- critical points $\begin{cases} 1. & \nabla f \text{ does not exist.} \\ 2. & \nabla f = \mathbf{0}. \end{cases}$

Q: Intuitively, why is this approach reasonable?

Exercise

(a) *Find the local extreme values, if any, of*

$$f(x, y) = x^2 + y^2 - 4y + 9$$

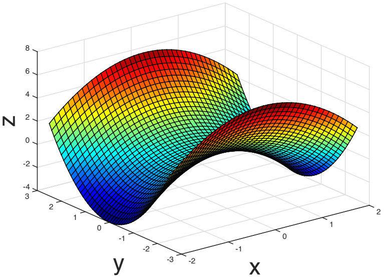
(b) *Determine the local extrema, if any, of*

$$f(x, y) = 1 - \sqrt[3]{x^2 + y^2}$$

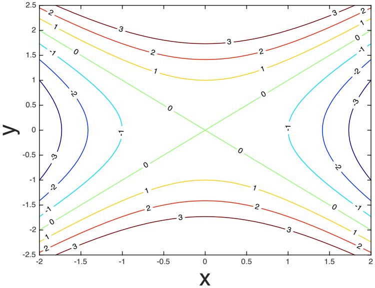
- As with differentiable function of a single variable, not every critical point leads a local extremum, it might be a point of inflection.
- A differentiable function of two variables might have a saddle point.

- Here is an example, the function has a saddle point at $(0,0)$.

$$f(x,y) = y^2 - x^2$$



$$y^2 - x^2 = k$$



Definition

We will say that a function has a saddle point P if there are two distinct vertical planes through P such that P in one of the planes is a local maximum and P in the other is a local minimum.

Exercise

Find the local extreme values, if any, of

$$z = y^2 - y^4 - x^2$$

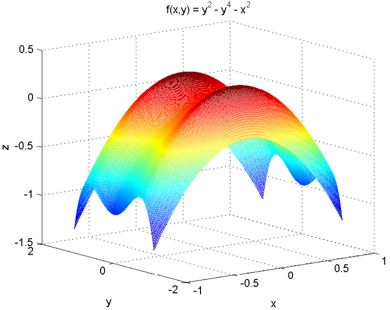
identify their nature by plotting a surface diagram and a contour of it.

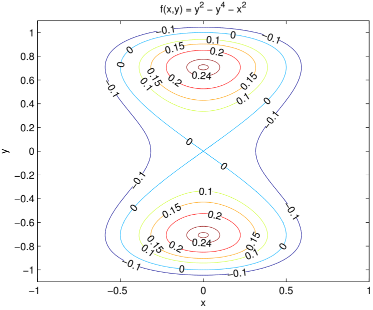
```
>> syms x y real; f_sym =  y^2 - y^4 -  x^2;
>> gradf=jacobian(f_sym,[x,y]);
>> sol = solve(gradf == 0);  % set the gradient to zero
>> sol.x % x coordinate
ans =    0    0    0
>> sol.y % x coordinate
ans =    0    2^(1/2)/2    -2^(1/2)/2

% f evaluated at those points
>> subs(f_sym, x,y, [sol.x],[sol.y])

ans =    0    1/4    1/4
```

- Based on the plots, we can conclude that





$$z = \frac{1}{4} \text{ at } (0, \frac{\sqrt{2}}{2}) \text{ and } (0, -\frac{\sqrt{2}}{2}) \text{ are local maxima of } z = y^2 - y^4 - x^2,$$

and the point $(0, 0, 0)$ is saddle point.

```
>> f = inline('y.^2-y.^4-x.^2','x','y'); [x, y] = meshgrid((-2:0.1:2),(-2.5:0.1:2.5));  z = f(x,y);
>> mesh(x,y,z); xlabel('x'); ylabel('y'); zlabel('z'); title('f(x,y) = y^2 - y^4 - x^2');

>> [k,h]=contour(x,y,z,[-0.1,0,0.1,0.15,0.2,0.24,0.249]);
>> clabel(k,h); xlabel('x'); ylabel('y'); title('f(x,y) = y^2 - y^4 - x^2');
```

Q: How can we determine the nature of a critical point of

$$z = f(x, y) \qquad \text{or} \qquad w = f(x, y, z)$$

- Recall for $y = f(x)$, we have the first and the second derivatives test.
- The first derivative test for a function of a single variable.

$$\begin{aligned} \text{positive} \quad f'(c) \quad \text{negative} &\implies \text{local maximum,} \\ \text{negative} \quad f'(c) \quad \text{positive} &\implies \text{local minimum,} \\ \text{No sign change} &\implies \text{not a local maximum or local minimum} \end{aligned}$$

- The second derivative test for at point c such $f'(c) = 0$.

$$\begin{aligned} f''(c) \quad \text{is positive} &\implies \text{local minimum at } c \\ f''(c) \quad \text{is negative} &\implies \text{local maximum at } c \\ f''(c) = 0 &\implies \text{inconclusive} \end{aligned}$$

Q: Why are we not going to have a first derivatives test in a similar way for

$$z = f(x, y) \qquad \text{or} \qquad w = f(x, y, z)$$

- However, we do have a second derivative test, which uses a special matrix.

Definition

If f is a function of n variables and if all second partial derivatives of f exist and are continuous over the domain of the function, then the Hessian matrix of f is

$$\mathbf{H} = \begin{bmatrix} f_{11} & f_{12} & f_{13} & \cdots & f_{1n} \\ f_{21} & f_{22} & f_{23} & \cdots & f_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ f_{n1} & f_{n2} & f_{n3} & \cdots & f_{nn} \end{bmatrix}$$

- For example, the Hessian matrix for $f(x, y) = x^2 y$ is

$$\mathbf{H} = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial x}(2xy) & \frac{\partial}{\partial y}(2xy) \\ \frac{\partial}{\partial x}\left(x^2\right) & \frac{\partial}{\partial y}\left(x^2\right) \end{bmatrix} = \begin{bmatrix} 2y & 2x \\ 2x & 0 \end{bmatrix}$$

- Consider the following function

$$g(x,y,z) = \left(x^2 + y^2 + z^2\right) e^x$$

it can be shown that $Q = (-2, 0, 0)$ is a critical point, that is,

$$\nabla g(-2,0,0) = \mathbf{0}$$

- The Hessian matrix at Q is given by

$$\mathbf{H} = \begin{bmatrix} g_{xx} & g_{xy} & g_{xz} \\ g_{yx} & g_{yy} & g_{yz} \\ g_{zx} & g_{zy} & g_{zz} \end{bmatrix} = e^x \begin{bmatrix} (x+2)^2 + y^2 + z^2 - 2 & 2y & 2z \\ 2y & 2 & 0 \\ 2z & 0 & 2 \end{bmatrix}$$

$$= e^{-2} \begin{bmatrix} -2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

- The nature of Q might be determined by using the eigenvalues of $\mathbf{H}$ at Q.

Definition

The scalar λ is called the eigenvalue of $\mathbf{A}_{n\times n}$ if there is a non-zero vector $\mathbf{x}$ of

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$$

Such non-zero vector $\mathbf{x}$ is known as the eigenvector corresponding to λ.

Theorem

The eigenvalue of a square matrix $\mathbf{A}$ satisfies the following polynomial equation

$$\det\left(\mathbf{A} - \lambda\mathbf{I}\right) = 0$$

where $\mathbf{I}$ is the identity matrix.

- To see why it is a polynomial, consider the the eigenvalue of

$$\mathbf{H} = \begin{bmatrix} -2e^{-2} & 0 & 0 \\ 0 & 2e^{-2} & 0 \\ 0 & 0 & 2e^{-2} \end{bmatrix}$$

- Find the determinant,

$$P(\lambda) = \det(\mathbf{H} - \lambda\mathbf{I}) = \det\left(\begin{bmatrix} -2e^{-2} & 0 & 0 \\ 0 & 2e^{-2} & 0 \\ 0 & 0 & 2e^{-2} \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right)$$

$$= \det\begin{bmatrix} -2e^{-2} - \lambda & 0 & 0 \\ 0 & 2e^{-2} - \lambda & 0 \\ 0 & 0 & 2e^{-2} - \lambda \end{bmatrix}$$

$$= -\left(\lambda + 2e^{-2}\right)\left(\lambda - 2e^{-2}\right)^2$$

- Therefore the eigenvalues of $\mathbf{H}$ of the function $g(x, y, z)$ at P are

$$-\left(\lambda + 2e^{-2}\right)\left(\lambda - 2e^{-2}\right)^2 = 0 \implies \begin{aligned} \lambda_1 &= -2e^{-2} \\ \lambda_2 &= 2e^{-2} \\ \lambda_3 &= 2e^{-2} \end{aligned}$$

- Back to our original question,

How to determine the nature of a critical point?

The second derivative test for a function of several variables

Suppose f is differentiable and $\nabla f = \mathbf{0}$ at a point P_0, then if

all the eigenvalues of $\mathbf{H}$ at P_0 are positive $\Longrightarrow$ local minimum

all the eigenvalues of $\mathbf{H}$ at P_0 are negative $\Longrightarrow$ local maximum

$\mathbf{H}$ at P_0 has both positive and negative eigenvalues $\Longrightarrow$ saddle point

One of the eigenvalues of $\mathbf{H}$ is zero $\Longrightarrow$ inconclusive

Q: What does the second derivative test say regarding

$$g(x, y, z) = (x^2 + y^2 + z^2)e^x \qquad \text{at} \qquad Q = (-2, 0, 0).$$

Exercise

Apply the second derivative test to determine the nature of the critical points of

$$z = y^2 - y^4 - x^2$$

```
>> syms x y real; f_sym =  y^2 - y^4 -  x^2;
>> gradf=jacobian(f_sym,[x,y]); % Finds the gradient
>> sol = solve(gradf == 0); % Sets the gradient to zero
>> hessianf = hessian(f_sym) % Finds the hessian matrix
hessianf =
[ -2,          0]
[  0, 2 - 12*y^2]
% Finds the eigenvalues of the hessian matrix
>> lambda = eig(hessianf);
>> n = numel (sol.x); % How many solutions
>> for i = 1 : n
>> disp([sol.x(i),sol.y(i)]);
>> disp(subs(lambda,[x,y],[sol.x(i),sol.y(i)]));
>> end
[ 0, 0]                 -2   2
[ 0,  2^(1/2)/2]        -2  -4
[ 0, -2^(1/2)/2]        -2  -4
```

- Recall the extreme-value theorem, EVT, states a function that is

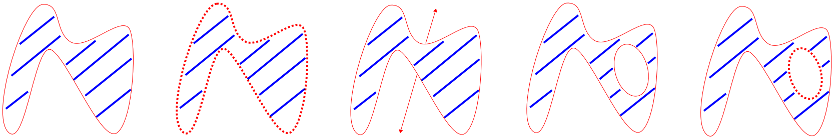
$$\underbrace{\text{continuous}}_{1.}\text{ throughout a }\underbrace{\text{closed and bounded}}_{2.}\text{ set }\mathcal{D},$$

  attains a global maximum value and a global minimum value at least once.
- For a function of a single variable, $y = f(x)$, the requirement 2. means a closed bounded interval or a finite union of closed bounded intervals

$$[a, b] \qquad \text{or} \qquad \bigcup_{i=1}^{n}[a_i, b_i]$$

- For a function of two or more variables, the requirement 2. means a region contains all of its boundary points and can be contained by some disk.

Q: Are the following sets closed and bounded?



- The EVT is sufficient but not necessary and it is an existence theorem.
- If extrema are guaranteed, then we can use the following steps to find them.

Procedures for finding global extrema

1. Find the local extreme values of f in the interior of the domain $\mathcal{D}$.
2. Find the local extreme values of f on boundary of the domain $\mathcal{D}$.
3. Compare values in step 1. and step 2., the largest of them is the global maximum, the smallest is the global minimum.

Exercise

Find the global maximum and minimum values of

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

on the triangular region in the first quadrant bounded by the lines

$$x = 0, \quad y = 0, \quad y = 9 - x$$

Exercise

A post-office accepts only rectangular boxes, of which the sum of whose length and perimeter of a cross-section does not exceed 108 cm. Find the dimensions of an acceptable box of largest volume.

```
>> syms y z real; V_sym = 108*y*z - 2*y^2*z - 2*y*z^2;
>> gradV=jacobian(V_sym,[y,z]); sol = solve(gradV == 0);
>> sol.y
ans = 0    54    0    18
>> sol.z
ans = 0    0    54    18

>> subs(V_sym,y,z, [sol.y],[sol.z])
ans = 0    0    0    11664

>> hessianV = hessian(V_sym);
>> lambda = eig(hessianV);
>> subs(lambda,[y,z],[sol.y(4),sol.z(4)])
ans = -108    -36
```