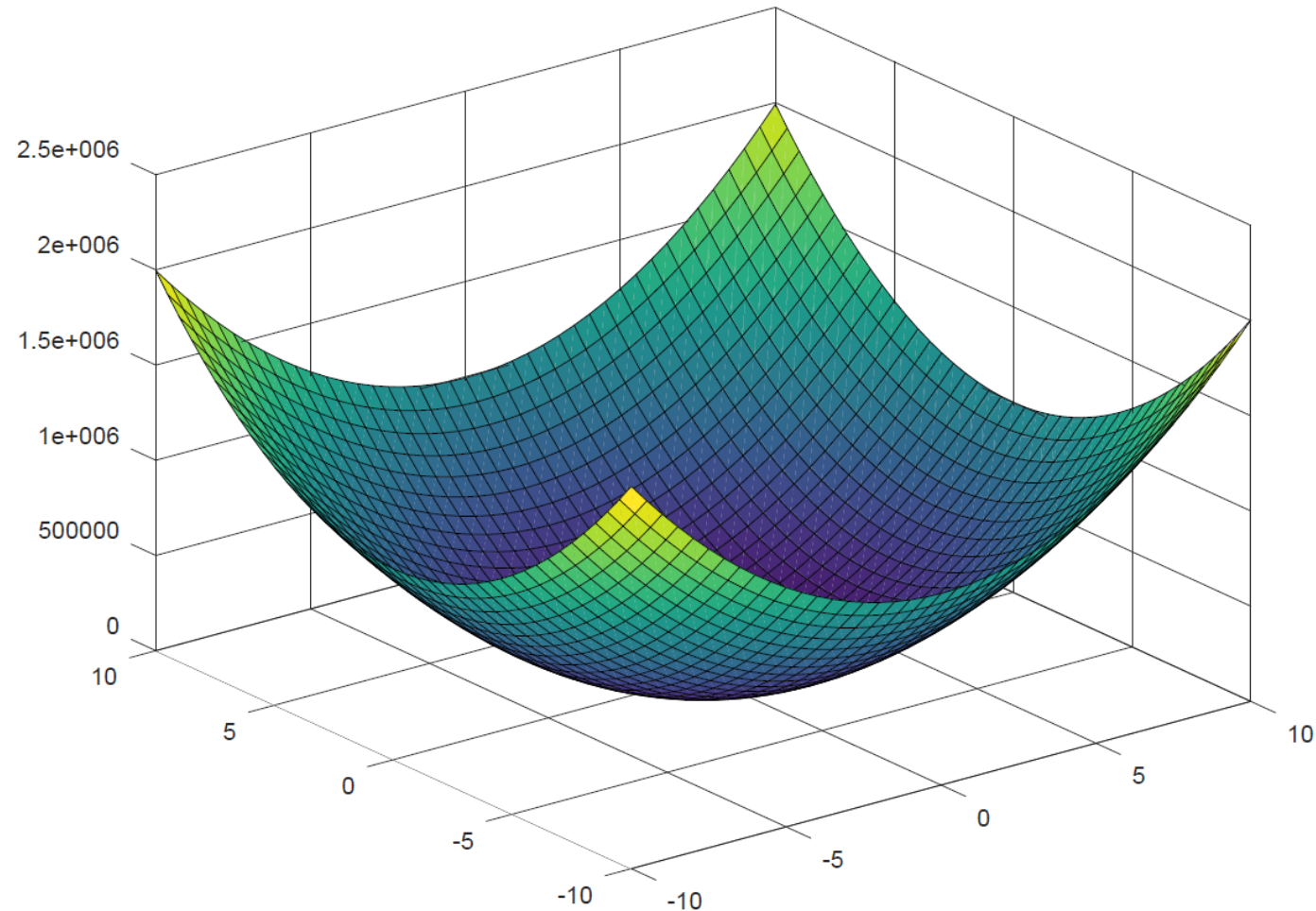


UP: Optimality Conditions



A. J. Conejo, R. Sioshansi 2016
THE OHIO STATE UNIVERSITY

What

1. First-Order Necessary Condition
2. Examples
3. Second-Order Sufficient Condition
4. Examples

Optimality Conditions

An unconstrained nonlinear optimization problem has the general form:

$$\min_{x \in \mathbb{R}^n} f(x),$$

where $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ is the objective being minimized and there are no constraints on the decision variables.

Optimality Conditions

$$\min_{x \in \mathbb{R}^n} f(x)$$

$$f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$$

First-Order Necessary Condition: FONC

Optimality Conditions: FONC

First-Order Necessary Condition (FONC): Consider an unconstrained nonlinear optimization problem of the form:

$$\min_{x \in \mathbb{R}^n} f(x).$$

If x^* is a local minimum, then $\nabla f(x^*) = 0$.

Optimality Conditions: FONC

$$\min_{x \in \mathbb{R}^n} f(x)$$
$$f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$$

If x^* is a local minimum, then $\nabla f(x^*) = 0$

FONC: proof

FONC: proof

Suppose x^* is a local minimum. Consider points, $x^* + d$, that are close to x^* (meaning that $\|d\|$ is close to zero). Then (Taylor):

$$f(x^* + d) \approx f(x^*) + d^\top \nabla f(x^*).$$

Because x^* is a local minimum, $f(x^* + d) \geq f(x^*)$.
Then:

$$f(x^*) + d^\top \nabla f(x^*) \geq f(x^*).$$

Finally:

$$d^\top \nabla f(x^*) \geq 0. \tag{1}$$

FONC: proof

Note that this must also apply for the point $x^* - d$. That is, $f(x^* - d) \geq f(x^*)$.

If we substitute the Taylor approximation for $f(x^* - d)$ into this inequality, this gives:

$$d^\top \nabla f(x^*) \leq 0 \quad (2)$$

FONC: proof

Combining (1) and (2) implies that we must have:

$$d^\top \nabla f(x^*) = 0.$$

Finally, note that we must have $d^\top \nabla f(x^*) = 0$ for any choice of d (so long as $\|d\|$ is close to zero).

The only way that this holds is if $\nabla f(x^*) = 0$.

Optimality Conditions: FONC

Points that satisfy the FONC (i.e., points that have a gradient equal to zero) are also known as

stationary points

Optimality Conditions: FONC

Square system of equations

It is important to note that the FONC always results in a system of n equations, where n is the number of decision variables in the optimization problem.

This is because the gradient is computed with respect to the decision variables, giving one first-order partial derivative for each decision variable.

Optimality Conditions: FONC

Not necessarily a
unique solution

The fact that the number of equations is equal to the number of variables does not necessarily mean that the FONC has a unique solution.

There could be multiple solutions or no solution.

Optimality Conditions: FONC

A important limitation of the FONC is stated below.

Although this condition eliminates non-stationary points that cannot be local minima, **it does not distinguish between local minima and local maxima.**

FONC: min vs. max

FONC: min vs. max

Consider the unconstrained problem:

$$\max_x f(x),$$

where the objective is an arbitrary function. To solve this problem, we convert it to a minimization of the form:

$$\min_x -f(x),$$

and search for stationary points:

$$\nabla(-f(x)) = -\nabla f(x) = 0.$$

FONC: min vs. max

Multiplying through by -1 , the FONC can be written as:

$$\nabla f(x) = 0,$$

which also is the FONC for the following problem:

$$\min_x f(x).$$

This means that the FONCs for finding local minima and local maxima of $f(x)$ are the same and the FONC cannot distinguish between the two.

FONC: Example

FONC: Example

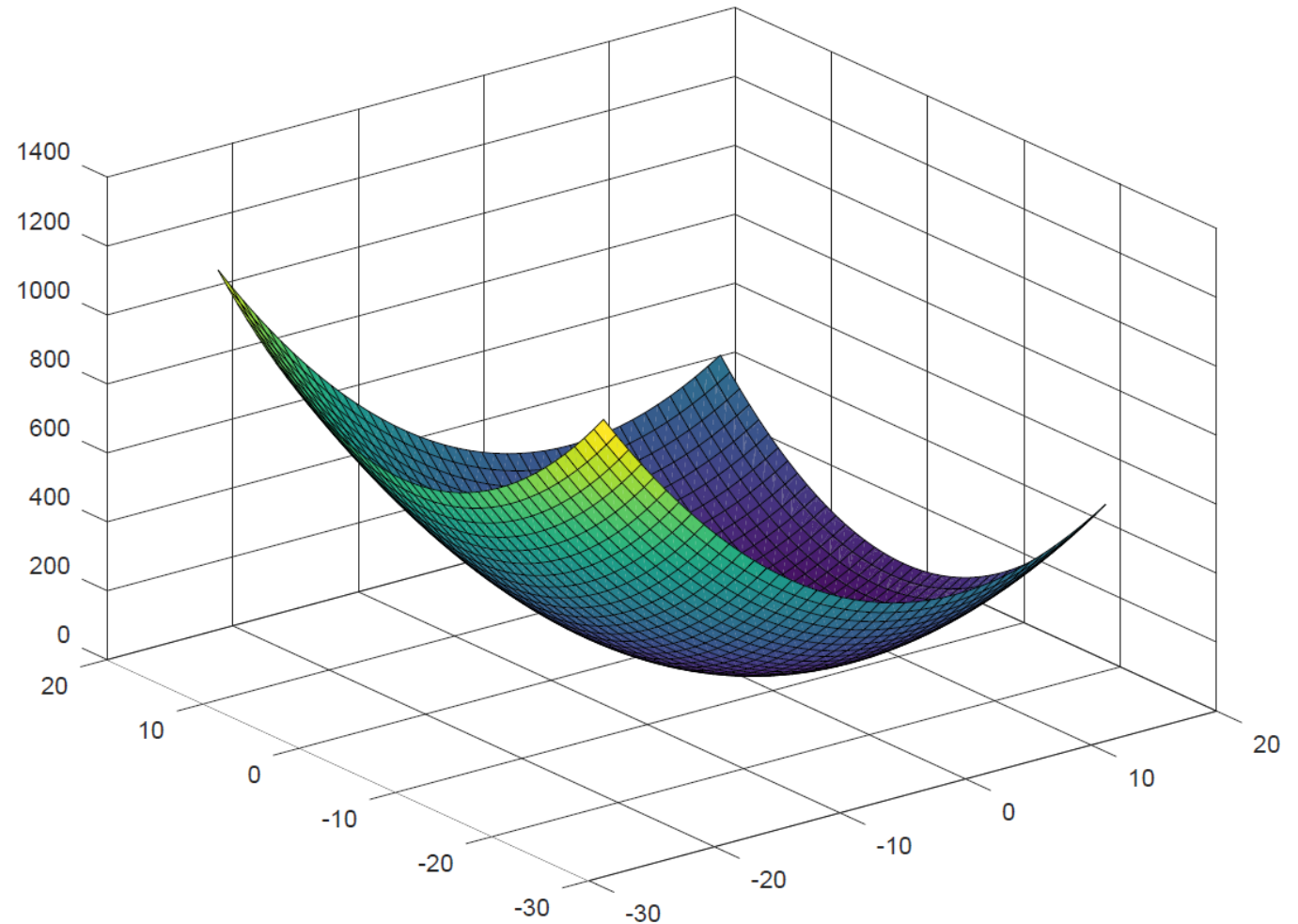
Consider the unconstrained problem:

$$\min_x f(x) = (x_1 - 3)^2 + (x_2 + 4)^2.$$

FONC: Example

$$\min_x f(x) = (x_1 - 3)^2 + (x_2 + 4)^2$$

```
clear  
x = -25:1:15;  
y = x;  
[X,Y] = meshgrid(x);  
F = (X.-3).^2+(Y.+4).^2;  
surf(X,Y,F)
```



FONC: Example

To find stationary points, we set the gradient of $f(x)$ equal to zero, which gives:

$$\nabla f(x) = \begin{pmatrix} 2(x_1 - 3) \\ 2(x_2 + 4) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

or:

$$x^* = \begin{pmatrix} 3 \\ -4 \end{pmatrix}.$$

$$\min_x f(x) = (x_1 - 3)^2 + (x_2 + 4)^2$$

FONC: Example

It is easy to confirm that this point is in fact a local and global minimum, because **the objective function is bounded below by zero.**

It we plug x^* into $f(x)$ we see that this point gives an objective-function value of zero, thus we have a local and global minimum.

FONC: Example

FONC: Example

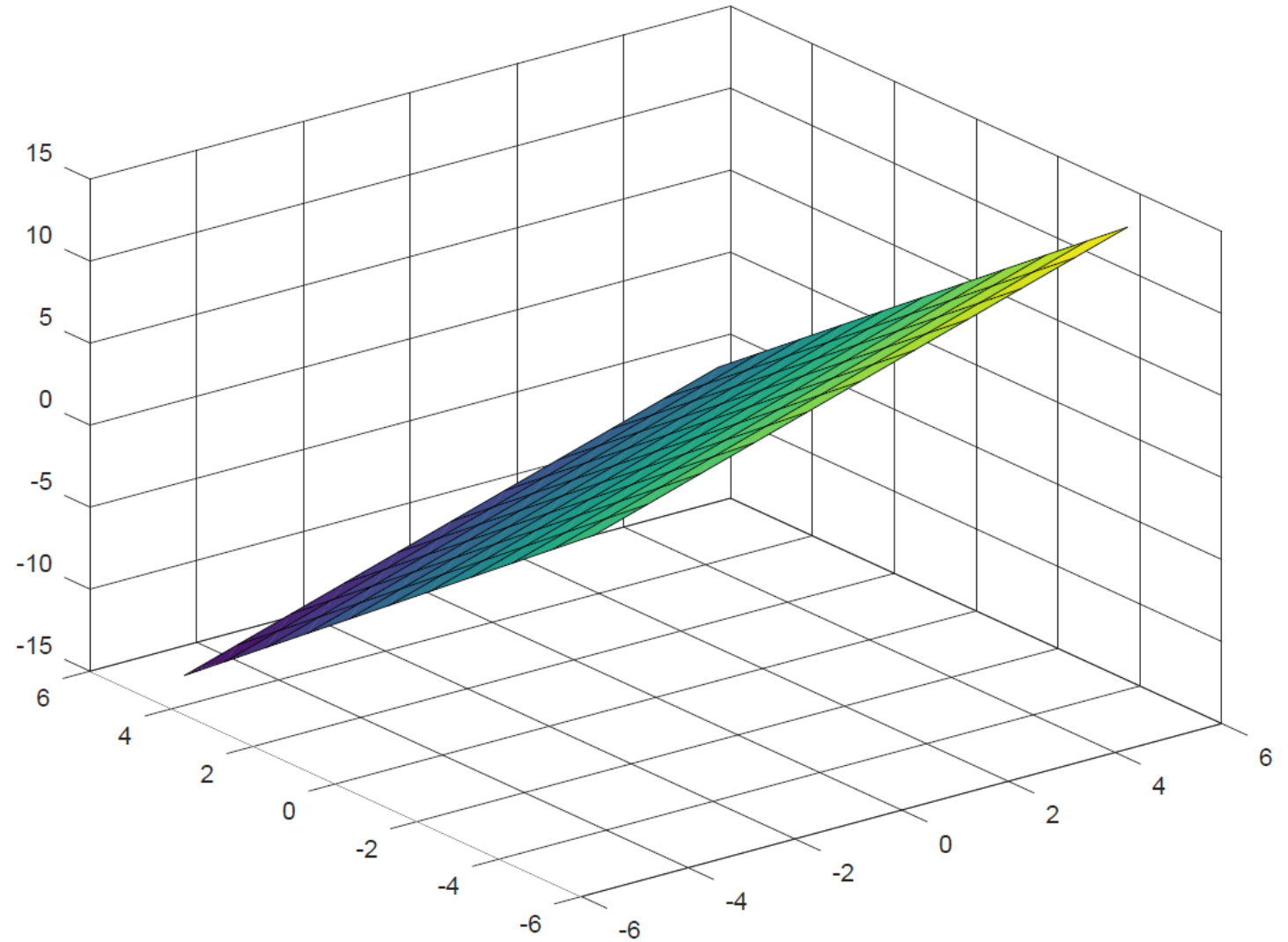
Consider the unconstrained problem:

$$\min_x f(x) = x_1 - 2x_2.$$

FONC: Example

$$\min_x f(x) = x_1 - 2x_2$$

```
clear
x = -5:1:5;
y = x;
[X,Y] = meshgrid(x);
F = (X)-2.*(Y);
surf(X,Y,F)
```



$$\min_x f(x) = x_1 - 2x_2$$

FONC: Example

The gradient of this objective function is:

$$\nabla f(x) = \begin{pmatrix} 1 \\ -2 \end{pmatrix},$$

which cannot be made to equal zero. Because this problem does not have any stationary points, it cannot have any local minima. In fact, this problem is unbounded. To see this, note that:

$$\lim_{x_1 \rightarrow -\infty, x_2 \rightarrow +\infty} f(x) = -\infty.$$

FONC: Example

FONC: Example

Consider the unconstrained problem:

$$\min_x f(x) = x^3 - x^2 - 4x - 6.$$

$$\min_x f(x) = x^3 - x^2 - 4x - 6$$

FONC: Example

To find stationary points, we set the gradient of $f(x)$ equal to zero, which gives:

$$\nabla f(x) = 3x^2 - 2x - 4 = 0,$$

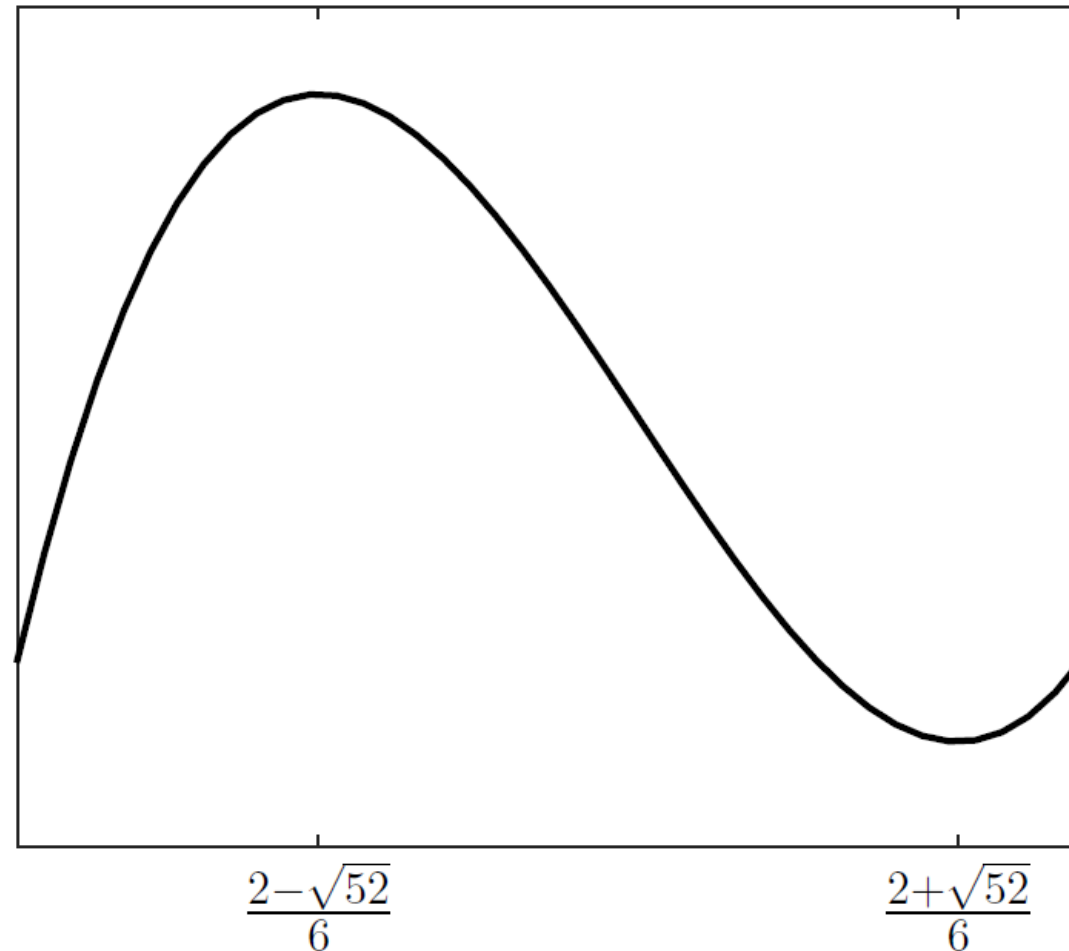
or:

$$x^* \in \left\{ \frac{2 - \sqrt{52}}{6}, \frac{2 + \sqrt{52}}{6} \right\}.$$

Both of these are stationary points and thus candidate local minima, based on the FONC.

FONC: Example

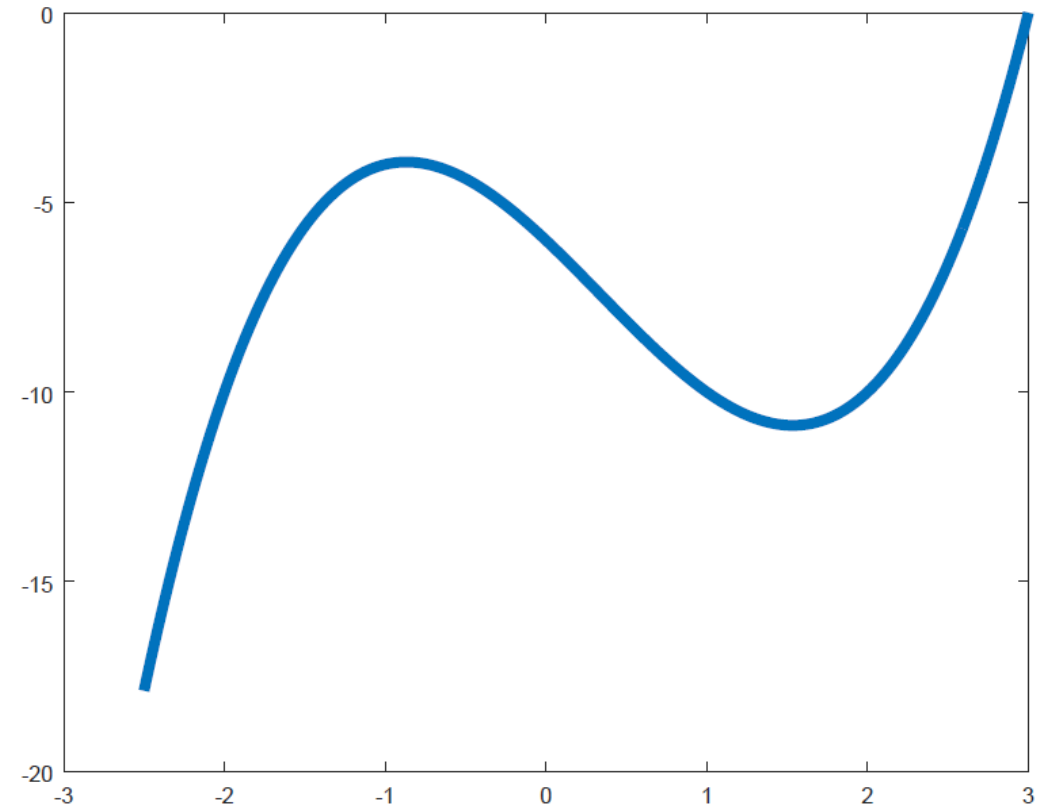
$$\min_x f(x) = x^3 - x^2 - 4x - 6$$



FONC: Example

$$\min_x f(x) = x^3 - x^2 - 4x - 6$$

```
home
x = -2.5:1/100:3;
y = power(x,3)-power(x,2)-4*x-6;
figure
plot(x,y,'LineWidth',5)
```



$$\min_x f(x) = x^3 - x^2 - 4x - 6$$

FONC: Example

The figure shows the objective function of this problem.

It is clear that only one of these two stationary points, $x^* = \frac{2+\sqrt{52}}{6}$, is a local minimum, whereas $\frac{2-\sqrt{52}}{6}$ is a local maximum.

$$\min_x f(x) = x^3 - x^2 - 4x - 6$$

FONC: Example

Moreover, this objective function is also unbounded. Although $x^* = \frac{2+\sqrt{52}}{6}$ is a local minimum, this particular problem does not have a global minimum because:

$$\lim_{x \rightarrow -\infty} f(x) = -\infty.$$

Second-Order Sufficient Condition: SOSOC

Second-Order Sufficient Condition: SOSOC

Consider an unconstrained nonlinear optimization problem of the form:

$$\min_{x \in \mathbb{R}^n} f(x).$$

If x^* satisfies $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*)$ is positive definite, then x^* is a local minimum.

Second-Order Sufficient Condition: SOSOC

$$\min_{x \in \mathbb{R}^n} f(x)$$
$$f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$$

If x^* satisfies $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*)$ is positive definite, then x^* is a local minimum.

SOSC: proof

Let λ be the smallest eigenvalue of $\nabla^2 f(x^*)$. If we consider a point $x^* + d$ that is close to x^* (i.e., by choosing a d with $\|d\|$ that is close to zero), the objective-function value at this point is approximately:

$$f(x^* + d) \approx f(x^*) + d^\top \nabla f(x^*) + \frac{1}{2} d^\top \nabla^2 f(x^*) d.$$

Because x^* is assumed to be stationary, we can rewrite this as:

$$f(x^* + d) - f(x^*) \approx \frac{1}{2} d^\top \nabla^2 f(x^*) d. \quad (1)$$

SOSC: proof

Check your algebra!

Using the Quadratic-Form Bound, we further have that:

$$\frac{1}{2}d^\top \nabla^2 f(x^*)d \geq \frac{1}{2}\lambda \|d\|^2 > 0, \quad (2)$$

where the last inequality follows from $\nabla^2 f(x^*)$ being positive definite, meaning that $\lambda > 0$.

SOSC: proof

Combining (1) and (2) gives:

$$f(x^* + d) - f(x^*) > 0,$$

meaning that x^* gives an objective-function value that is strictly better than any point that is close to it.

Thus, x^* is a local minimum.

SOSC: Example

SOSC: Example

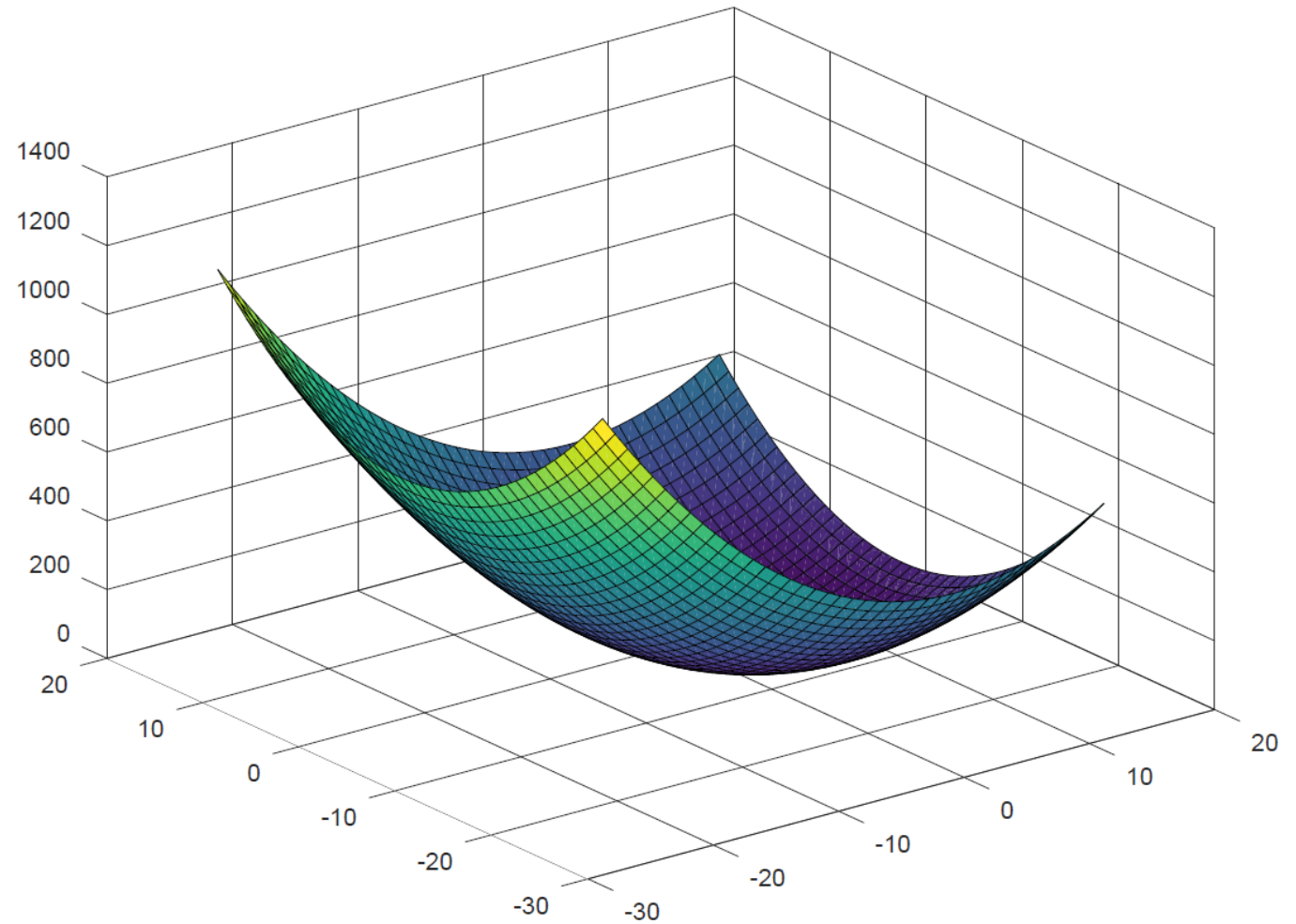
Consider the unconstrained optimization problem:

$$\min_x f(x) = (x_1 - 3)^2 + (x_2 + 4)^2.$$

SOSC: Example

$$\min_x f(x) = (x_1 - 3)^2 + (x_2 + 4)^2$$

```
clear
x = -25:1:15;
y = x;
[X,Y] = meshgrid(x);
F = (X.-3).^2+(Y.+4).^2;
surf(X,Y,F)
```



SOSC: Example

$$\min_x f(x) = (x_1 - 3)^2 + (x_2 + 4)^2$$

$$\nabla^2 f(x^*) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

This problem has a single stationary point:

$$x^* = \begin{pmatrix} 3 \\ -4 \end{pmatrix},$$

and the Hessian of the objective function is positive definite at this point.

Because this point satisfies the SOSC, it is guaranteed to be a local minimum.

SOSC: Example

SOSC: Example

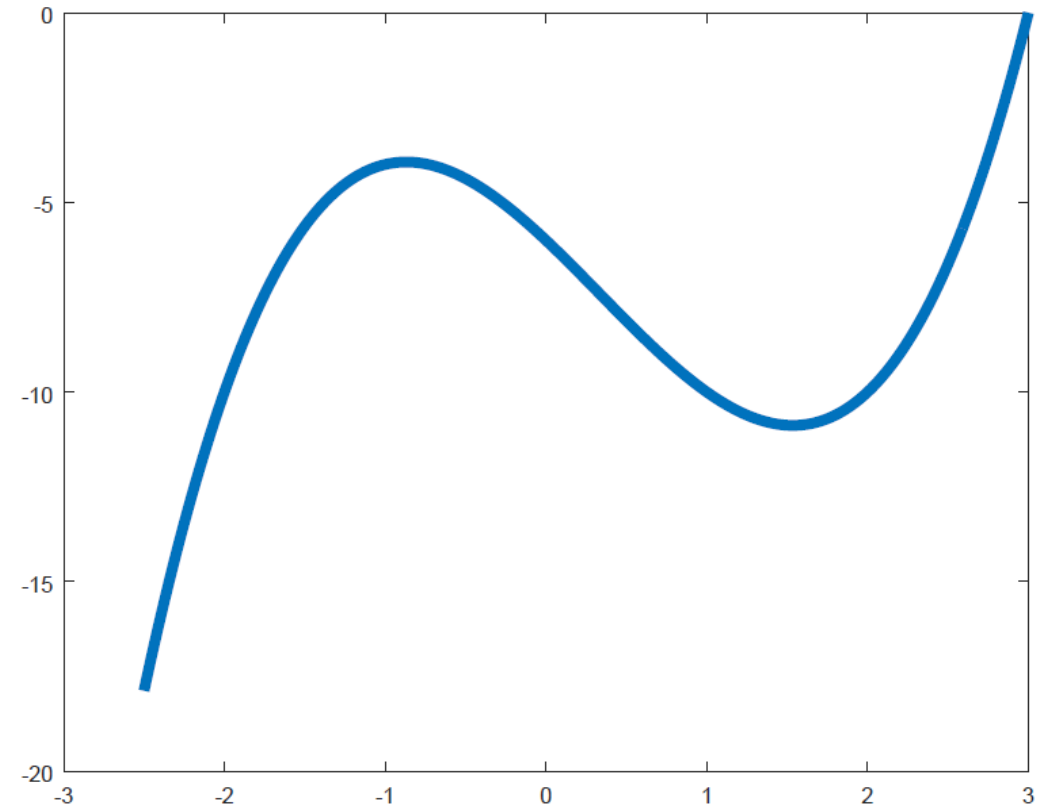
Consider the unconstrained optimization problem:

$$\min_x f(x) = x^3 - x^2 - 4x - 6.$$

$$\min_x f(x) = x^3 - x^2 - 4x - 6$$

SOSC: Example

```
home  
x = -2.5:1/100:3;  
y = power(x,3)-power(x,2)-4*x-6;  
figure  
plot(x,y,'LineWidth',5)
```



$$\min_x f(x) = x^3 - x^2 - 4x - 6$$

SOSC: Example

$$\nabla^2 f(x) = 6x - 2$$

This problem has two stationary points:

$$x^* \in \left\{ \frac{2 - \sqrt{52}}{6}, \frac{2 + \sqrt{52}}{6} \right\},$$

and the Hessian of the objective is positive definite at $(2 + \sqrt{52})/6$.

Because this point satisfies the SOSC, it is guaranteed to be a local minimum.

SOSC Limitation

SOSC Limitation

A limitation of the SOSC is that it is **not a necessary condition** for a point to be a local minimum. This means that there can be points that are local minima, yet do not satisfy the SOSC.

We demonstrate this with the following example.

SOSC: Example

SOSC: Example

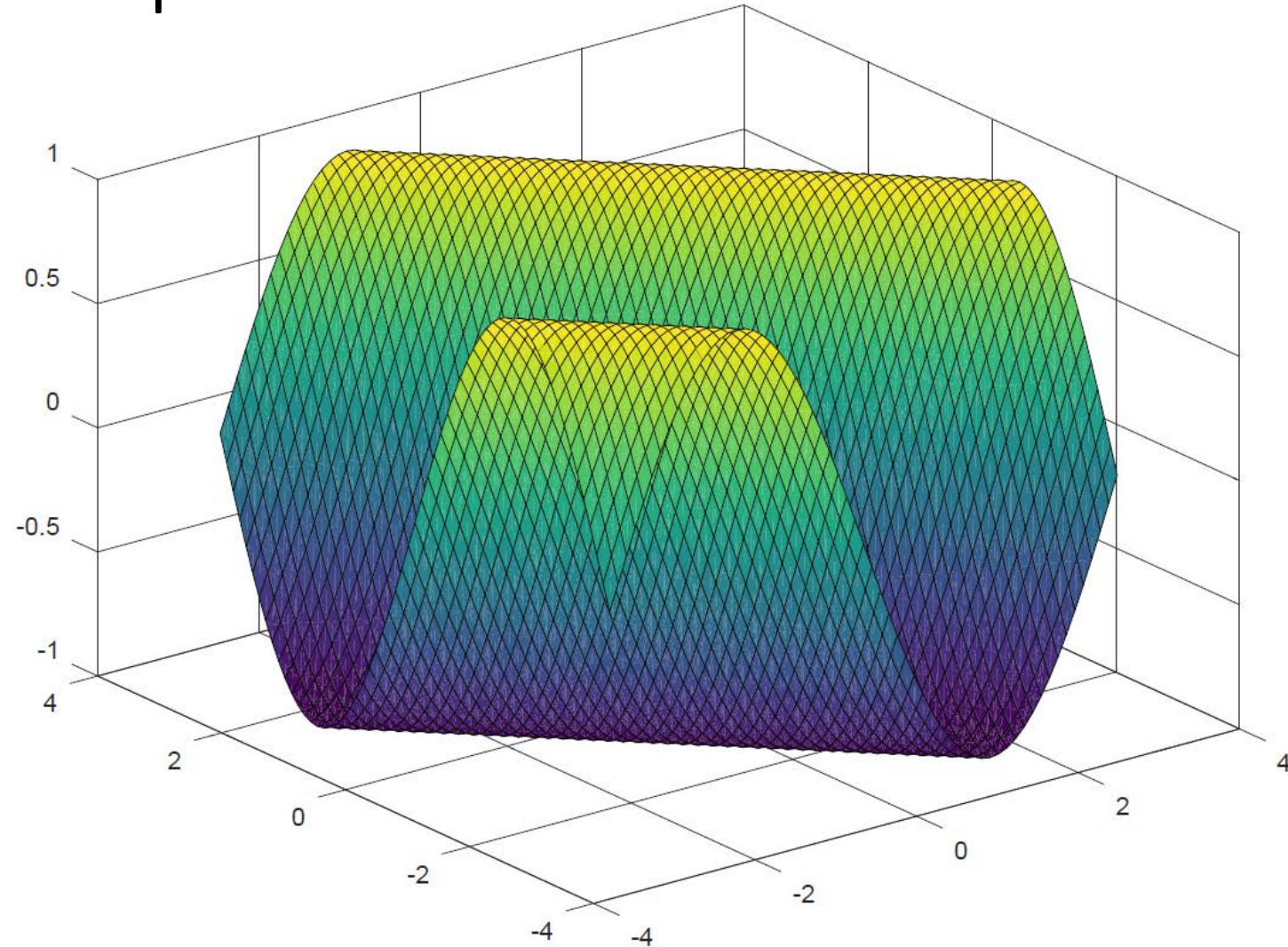
Consider the unconstrained optimization problem:

$$\min_x f(x) = \sin(x_1 + x_2)$$

SOSC: Example

```
home
x = -2*pi:pi/100:2*pi;
y = -2*pi:pi/100:2*pi;
[X,Y] = meshgrid(x,y);
Z = sin(X+Y);
figure
surf(X,Y,Z)
```

SOSC: Example



$$\min_x f(x) = \sin(x_1 + x_2)$$

SOSC: Example

The FONC is:

$$\nabla f(x) = \begin{pmatrix} \cos(x_1 + x_2) \\ \cos(x_1 + x_2) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

$$\min_x f(x) = \sin(x_1 + x_2)$$

SOSC: Example

Points of the form:

$$x^* = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

where:

$$x_1 + x_2 \in \left\{ \dots, -\frac{5\pi}{2}, -\frac{\pi}{2}, \frac{3\pi}{2}, \frac{7\pi}{2}, \dots \right\},$$

satisfy the FONC.

$$\min_x f(x) = \sin(x_1 + x_2)$$

SOSC: Example

The Hessian of the objective function is:

$$\nabla^2 f(x) = \begin{bmatrix} -\sin(x_1 + x_2) & -\sin(x_1 + x_2) \\ -\sin(x_1 + x_2) & -\sin(x_1 + x_2) \end{bmatrix}.$$

$$\min_x f(x) = \sin(x_1 + x_2)$$

SOSC: Example

However the Hessian is only positive semidefinite (as opposed to positive definite).

Thus, these points do not satisfy the SOSC and are not guaranteed to be local minima on the basis of that optimality condition.

$$\min_x f(x) = \sin(x_1 + x_2)$$

SOSC: Example

However, because the objective function attains its lower bound of -1 at these points, they must be local and global minima.

Thus, this problem has local minima that do not satisfy the SOSC.



SOSC: Example

SOSC: Example

Consider the unconstrained optimization problem:

$$\min_x f(x) = x^3.$$

$$\min_x f(x) = x^3$$

SOSC: Example

The FONC is $\nabla f(x) = 3x^2 = 0$, which gives $x^* = 0$ as the unique stationary point.

We further have that $\nabla^2 f(x) = 6x$ and substituting $x^* = 0$ into the Hessian gives a value of zero, meaning that it is positive semidefinite.

$$\min_x f(x) = x^3$$

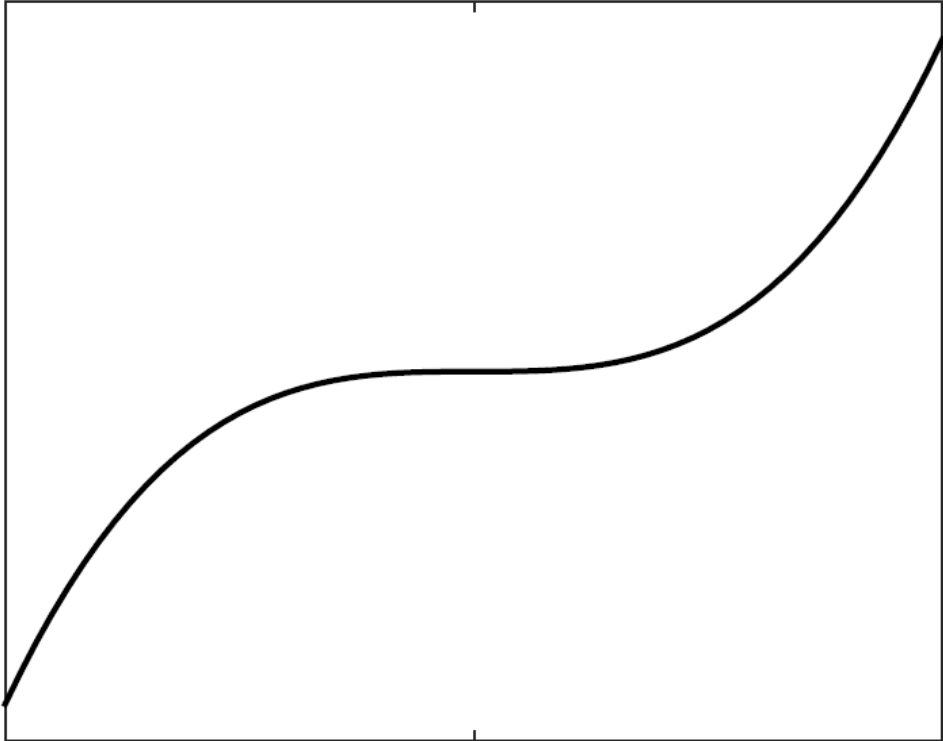
SOSC: Example

However, visual inspection of the objective function, which is shown in the figure below, shows that this point is not a local minimum.

Instead it is a point of inflection or a saddle point.

$$\min_x f(x) = x^3$$

SOSC: Example



$$x^* = 0$$

Solving the FONC



**Solving the
FONC**

Solving the FONC

$$\min_{x_1, x_2} z = (x_1 - 1)^4 + (x_2 - 2)^4$$

Solving the FONC

$$\min_{x_1, x_2} z = (x_1 - 1)^4 + (x_2 - 2)^4$$

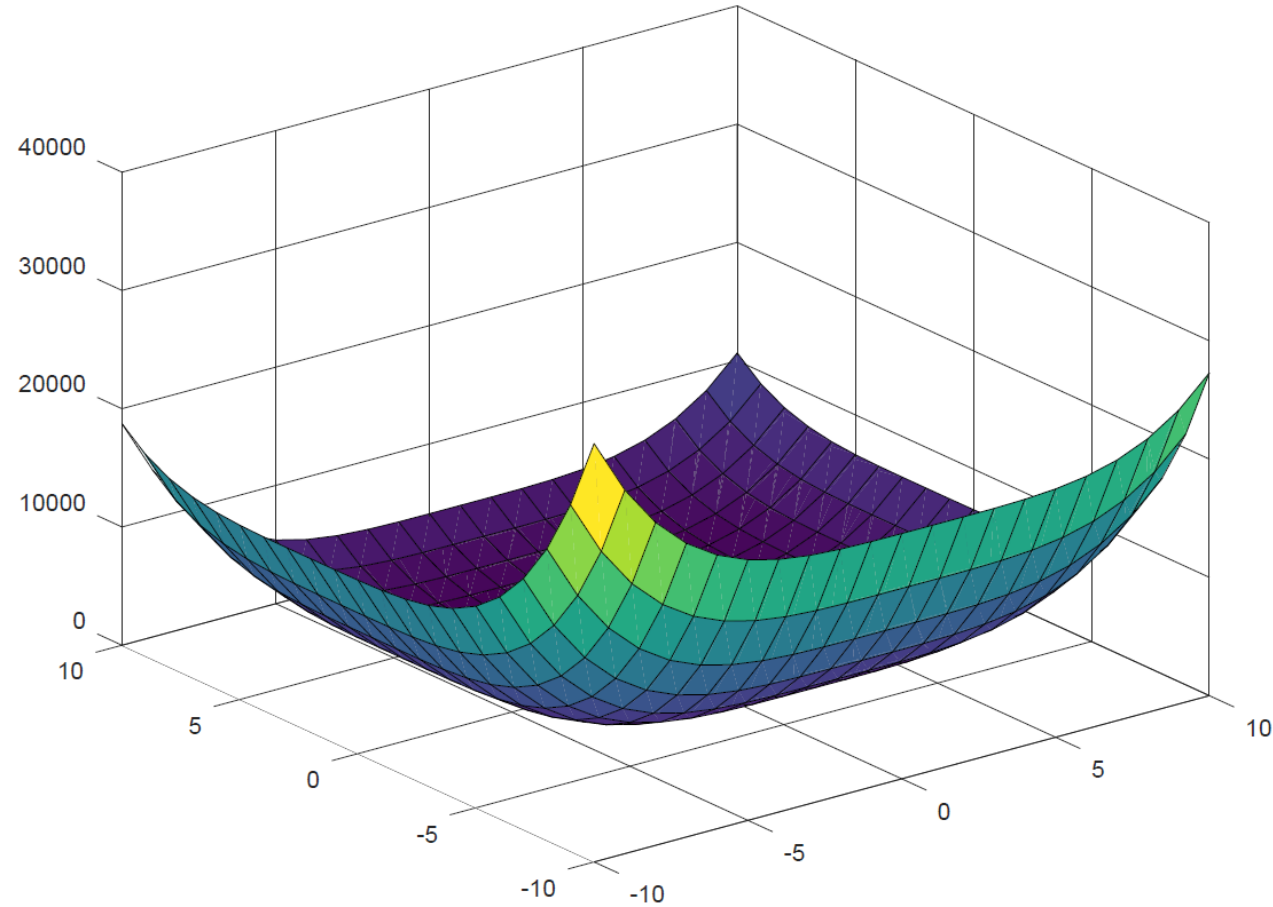
Inspection!

```
clear
x1 = -10:1:10;
x2 = x1;
[X1,X2] = meshgrid(x1,x2);
F = (X1.-1).^4+(X2.-2).^4;
figure
surf(X1,X2,F)
```

Solving the FONC

$$\min_{x_1, x_2} z = (x_1 - 1)^4 + (x_2 - 2)^4$$

Inspection!



Solving the FONC

$$\min_{x_1, x_2} z = (x_1 - 1)^4 + (x_2 - 2)^4$$



```
variable z;  
variable x1, x2;  
x1.l=1; x2.l=1;  
equation fo;  
fo.. z =e= power((x1-1),4)+power((x2-2),4);  
model solsin /all/;  
solve solsin minimizing z using nlp;
```

Solving the FONC

$$\min_{x_1, x_2} z = (x_1 - 1)^4 + (x_2 - 2)^4$$



LOWER	LEVEL	UPPER	MARGINAL		
----	VAR z	-INF	1.454E-11	+INF	.
----	VAR x1	-INF	1.000	+INF	EPS
----	VAR x2	-INF	1.998	+INF	-2.979E-8

Solving the FONC

$$\min_{x_1, x_2} z = (x_1 - 1)^4 + (x_2 - 2)^4$$

Octave!

m-file solvingfonc1.m

```
clear
fun = @fonc1;
x0 = [1,1];
x = fsolve(fun,x0)
```

m-file fonc1.m

```
function F = fonc1(x)
%
F(1)=4*(x(1)-1)^3;
F(2)=4*(x(2)-2)^3;
```


Solving the FONC

$$\min_{x_1, x_2} z = (x_1 - 1)^4 + (x_2 - 2)^4$$

The screenshot displays the Octave software interface. The main window is titled "Octave" and contains several panels:

- File Browser:** Shows the current directory as `5 to 2.05)\Class material\Octave\UC OC\solving`. It lists two files: `fonc1.m` and `solvingfonc1.m`.
- Command Window:** Shows the command `>> solvingfonc1` and the resulting output:

```
X =  
1.0000 1.9956  
>>
```
- Workspace:** A table showing the current workspace variables:

Name	Class	Dimension	Value
fun	function...	1x1	...
x	double	1x2	[1, 1.9956]
x0	double	1x2	[1, 1]
- Command History:** Shows the sequence of commands: `exit`, `# Octave 4.2.0, Mon Jan 16 11:53:56 2`, `home`, and `solvingfonc1`.

A large red oval with the text "Octave!" is overlaid on the right side of the Command Window.

This is it!