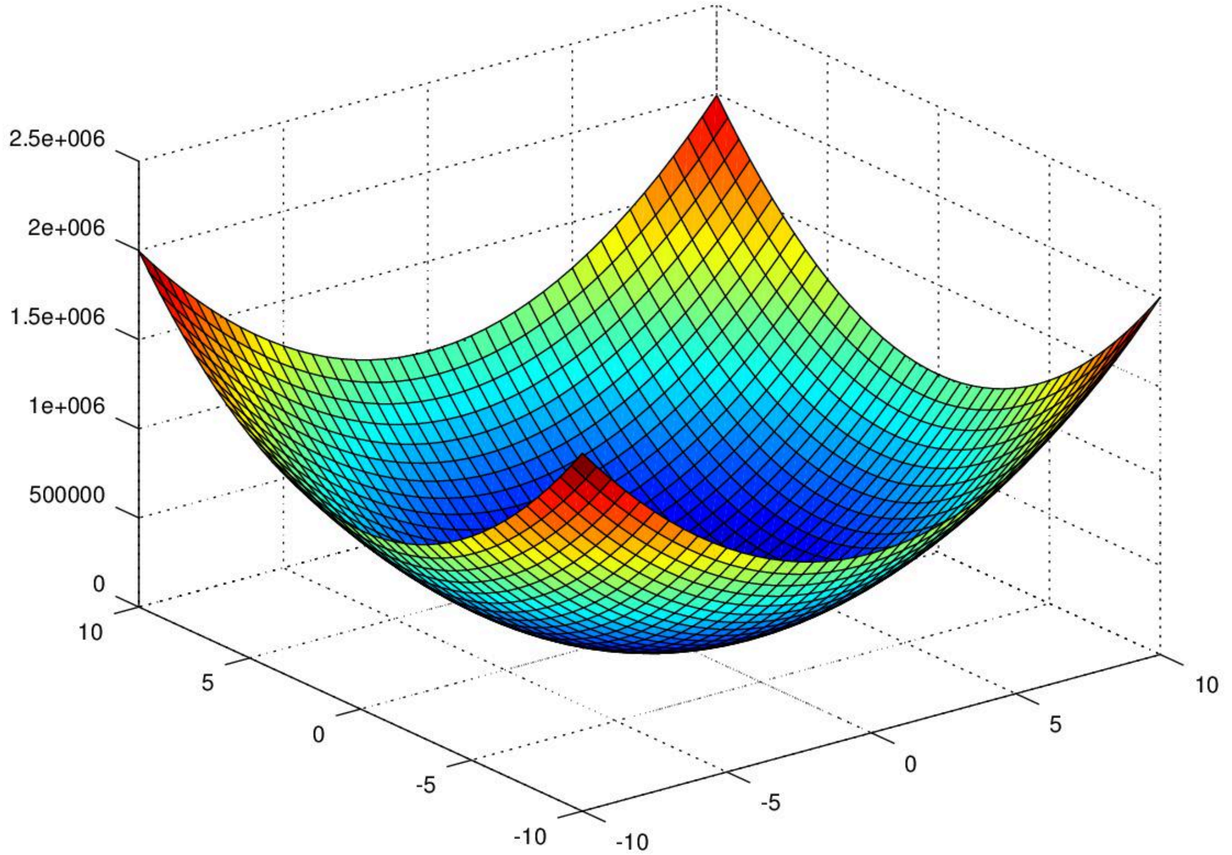
ECP: Optimality Conditions



A. J. Conejo, R. Sioshansi 2016

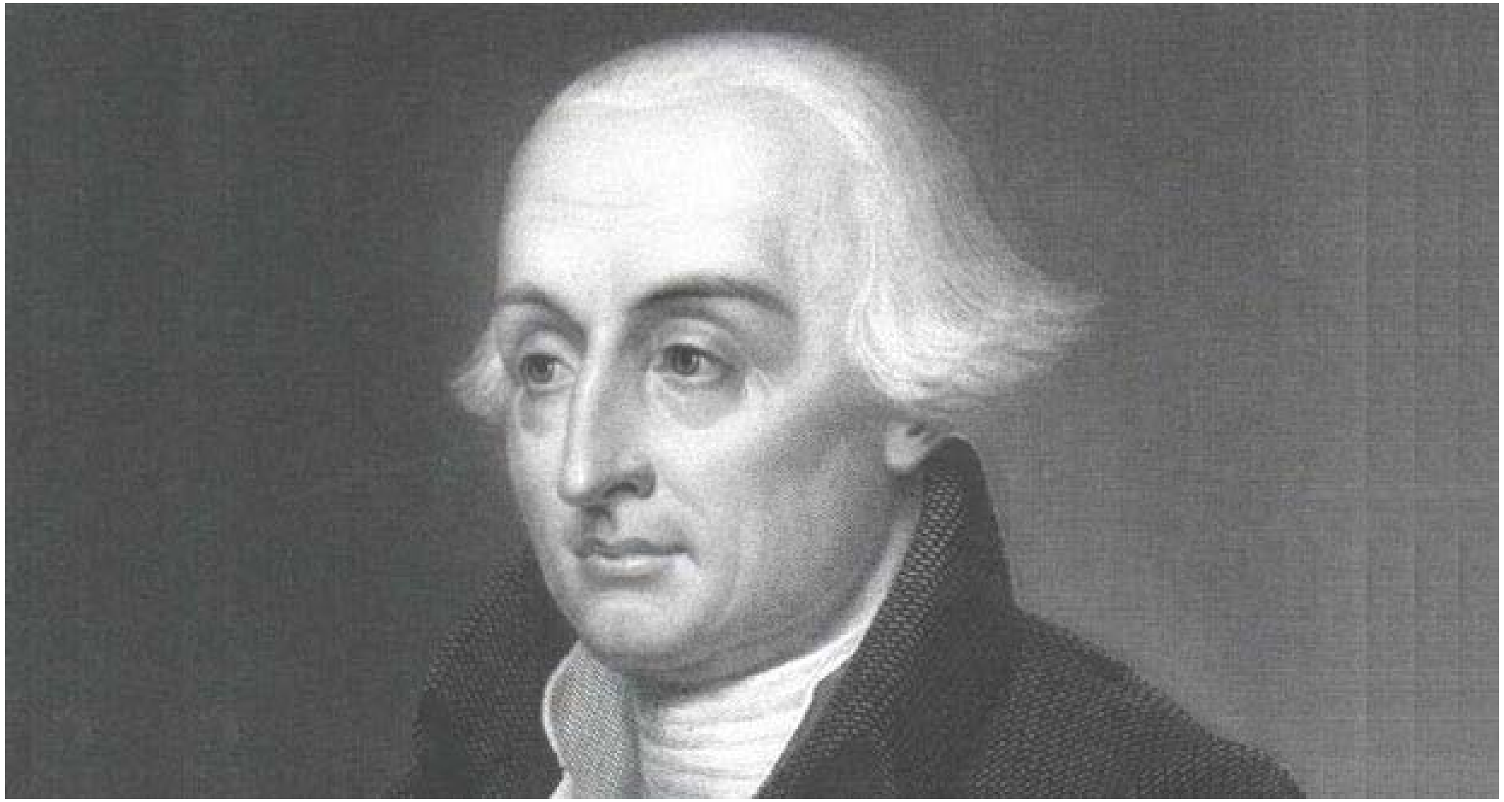
THE OHIO STATE UNIVERSITY

What

Equality-Constrained Nonlinear Optimization Problems

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

Joseph Louis Lagrange



FONC

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

$$\text{s.t.} \quad h_1(x) = 0$$

$$h_2(x) = 0$$

$$\vdots$$

$$h_m(x) = 0$$

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

$$h_i(x) : \mathbb{R}^n \to \mathbb{R}, i = 1, \ldots, m$$

FONC

Consider an equality-constrained nonlinear optimization problem of the form:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \ & f(x) \\ \text{s.t.} \ & h_1(x) = 0 \\ & h_2(x) = 0 \\ & \vdots \\ & h_m(x) = 0. \end{aligned}$$

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

$$h_i(x) : \mathbb{R}^n \to \mathbb{R}, i = 1, \ldots, m$$

If x^* is a local minimum of this problem, then there exist m **Lagrange multipliers**, $\lambda_1^*, \lambda_2^*, \ldots, \lambda_m^*$, such that:

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla h_i(x^*) = 0.$$

FONC

$$\min_{x \in \mathbb{R}^n} f(x)$$
$$\text{s.t.} \quad h_1(x) = 0$$
$$h_2(x) = 0$$
$$\vdots$$
$$h_m(x) = 0$$

$$f(x) : \mathbb{R}^n \to \mathbb{R}$$
$$h_i(x) : \mathbb{R}^n \to \mathbb{R}, i = 1, \ldots, m$$

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla h_i(x^*) = 0$$

FONC

The FONC for an equality-constrained problem require us to solve not only for values of the decisions variables in the original problem (i.e., the x's) but also for an additional set of m Lagrange multipliers.

FONC

Note that the number of Lagrange multipliers is always equal to the number of equality constraints in the original problem.

FONC: Example

FONC: Example

Consider the equality-constrained problem:

$$\min_{x} \; f(x) = 4x_1^2 + 3x_2^2 + 2x_1x_2 + 4x_1 + 6x_2 + 3$$

$$\text{s.t.} \quad h_1(x) = x_1 - 2x_2 - 1 = 0$$

$$h_2(x) = x_1^2 + x_2^2 - 1 = 0.$$

FONC: Example

$$\min_x \ f(x) = 4x_1^2 + 3x_2^2 + 2x_1x_2 + 4x_1 + 6x_2 + 3$$
$$\text{s.t.} \ \ h_1(x) = x_1 - 2x_2 - 1 = 0$$
$$h_2(x) = x_1^2 + x_2^2 - 1 = 0.$$

To apply the FONC, we define two Lagrange multipliers, λ_1 and λ_2, which are associated with the two constraints. The FONC is then:

$$\nabla f(x^*) + \sum_{i=1}^{2} \lambda_i^* \nabla h_i(x^*) = 0,$$

or:

$$\begin{pmatrix} 8x_1^* + 2x_2^* + 4 \\ 6x_2^* + 2x_1^* + 6 \end{pmatrix} + \lambda_1^* \begin{pmatrix} 1 \\ -2 \end{pmatrix} + \lambda_2^* \begin{pmatrix} 2x_1^* \\ 2x_2^* \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

FONC: Example

Note that this is a system of two equations with four unknowns—the two original decision variables (x_1 and x_2) and the two Lagrange multipliers (λ_1 and λ_2).

We do have two additional conditions that the x's must satisfy, which are the original constraints of the problem.

$$\min_{x} \; f(x) = 4x_1^2 + 3x_2^2 + 2x_1 x_2 + 4x_1 + 6x_2 + 3$$
$$\text{s.t.} \;\; h_1(x) = x_1 - 2x_2 - 1 = 0$$
$$h_2(x) = x_1^2 + x_2^2 - 1 = 0.$$

FONC: Example

If we add these two constraints, we now have the following system of four equations with four unknowns:

$$8x_1^* + 2x_2^* + 4 + \lambda_1^* + 2\lambda_2^* x_1^* = 0$$

$$6x_2^* + 2x_1^* + 6 - 2\lambda_1^* + 2\lambda_2^* x_2^* = 0$$

$$x_1^* - 2x_2^* - 1 = 0$$

$$x_1^{*2} + x_2^{*2} - 1 = 0.$$

FONC: Example

This system of equations has two solutions:

$$(x_1^*, x_2^*, \lambda_1^*, \lambda_2^*) = (1, 0, 4, -8),$$

and:

$$(x_1^*, x_2^*, \lambda_1^*, \lambda_2^*) = (-3/5, -4/5, 24/25, -6/5).$$

FONC: Example

Because these are the only two values of x and λ that satisfy the constraints of the problem and the FONC, the candidate values of x that can be local minima are:

$$\begin{pmatrix} x_1^* \\ x_2^* \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

and:

$$\begin{pmatrix} x_1^* \\ x_2^* \end{pmatrix} = \begin{pmatrix} -3/5 \\ -4/5 \end{pmatrix}.$$

FONC: Example

We know that this problem is bounded, because the feasible region is bounded and the objective function does not asymptote. Thus, we know one of these two candidate points must be a global minimum. If we plug these values into the objective function we have:

$$f\begin{pmatrix} 1 \\ 0 \end{pmatrix} = 11,$$

and:

$$f\begin{pmatrix} -3/5 \\ -4/5 \end{pmatrix} = \frac{3}{25}.$$

FONC: Example

Because it gives a smaller objective-function value, we know that:

$$\begin{pmatrix} x_1^* \\ x_2^* \end{pmatrix} = \begin{pmatrix} -3/5 \\ -4/5 \end{pmatrix},$$

is the global minimum of this problem.

FONC: Example

FONC: Example

Consider the equality-constrained problem:

$$\min_{x,y,z} \; f(x) = (x-3)^2 - 10 + (2y-4)^2 - 14 + (z-6)^2 - 6$$

$$\text{s.t.} \;\; h_1(x) = x + y + z - 10 = 0.$$

FONC: Example

$$\min_{x,y,z} f(x) = (x-3)^2 - 10 + (2y-4)^2 - 14 + (z-6)^2 - 6$$

$$\text{s.t. } h_1(x) = x + y + z - 10 = 0.$$

To solve this problem we introduce one Lagrange multiplier, λ_1. The FONC and the original constraint of the problem give us the following system of equations:

$$2(x-3) + \lambda_1 = 0$$

$$4(2y-4) + \lambda_1 = 0$$

$$2(z-6) + \lambda_1 = 0$$

$$x + y + z - 10 = 0.$$

FONC: Example

$$\min_{x,y,z} f(x) = (x-3)^2 - 10 + (2y-4)^2 - 14 + (z-6)^2 - 6$$
$$\text{s.t.} \quad h_1(x) = x + y + z - 10 = 0.$$

The one solution to this system of equations is:

$$(x^*, y^*, z^*, \lambda_1^*) = (23/9, 17/9, 50/9, 8/9).$$

$$\min_{x,y,z} \; f(x) = (x-3)^2 - 10 + (2y-4)^2 - 14 + (z-6)^2 - 6$$
$$\text{s.t.} \;\; h_1(x) = x + y + z - 10 = 0.$$

FONC: Example

The constraint of this problem is linear and the Hessian of the objective function is:

$$\nabla^2 f(x,y,z) = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 2 \end{bmatrix},$$

which is positive definite, meaning that the objective is convex. Thus, the solution to the FONC that we have found is guaranteed to be a global minimum.

FONC

FONC

Just as in the unconstrained case, the FONC give us candidate solutions that could be local minima.

Moreover, points that do not satisfy the FONC *cannot* be local minima.

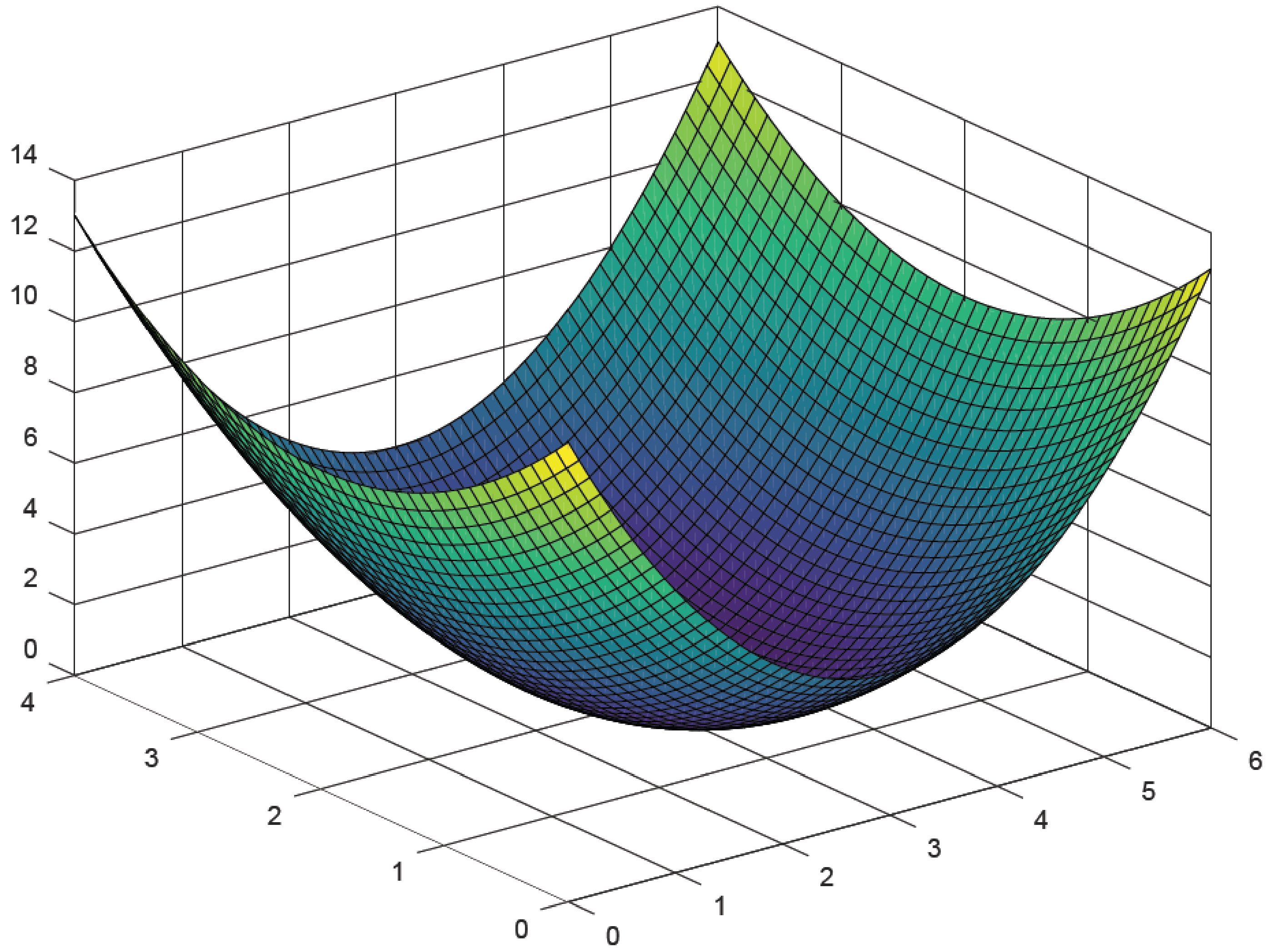
Thus, the FONC typically eliminate many possible points from further consideration.

Nevertheless, the FONC cannot necessarily distinguish between local minima, local maxima, and saddle points.

Surface

```
home
x = 0:0.1:6;
y = 0:0.1:4;
[X,Y] = meshgrid(x,y);
Z = (X.-3).^2+(Y.-2).^2;
figure
surf(X,Y,Z)
```

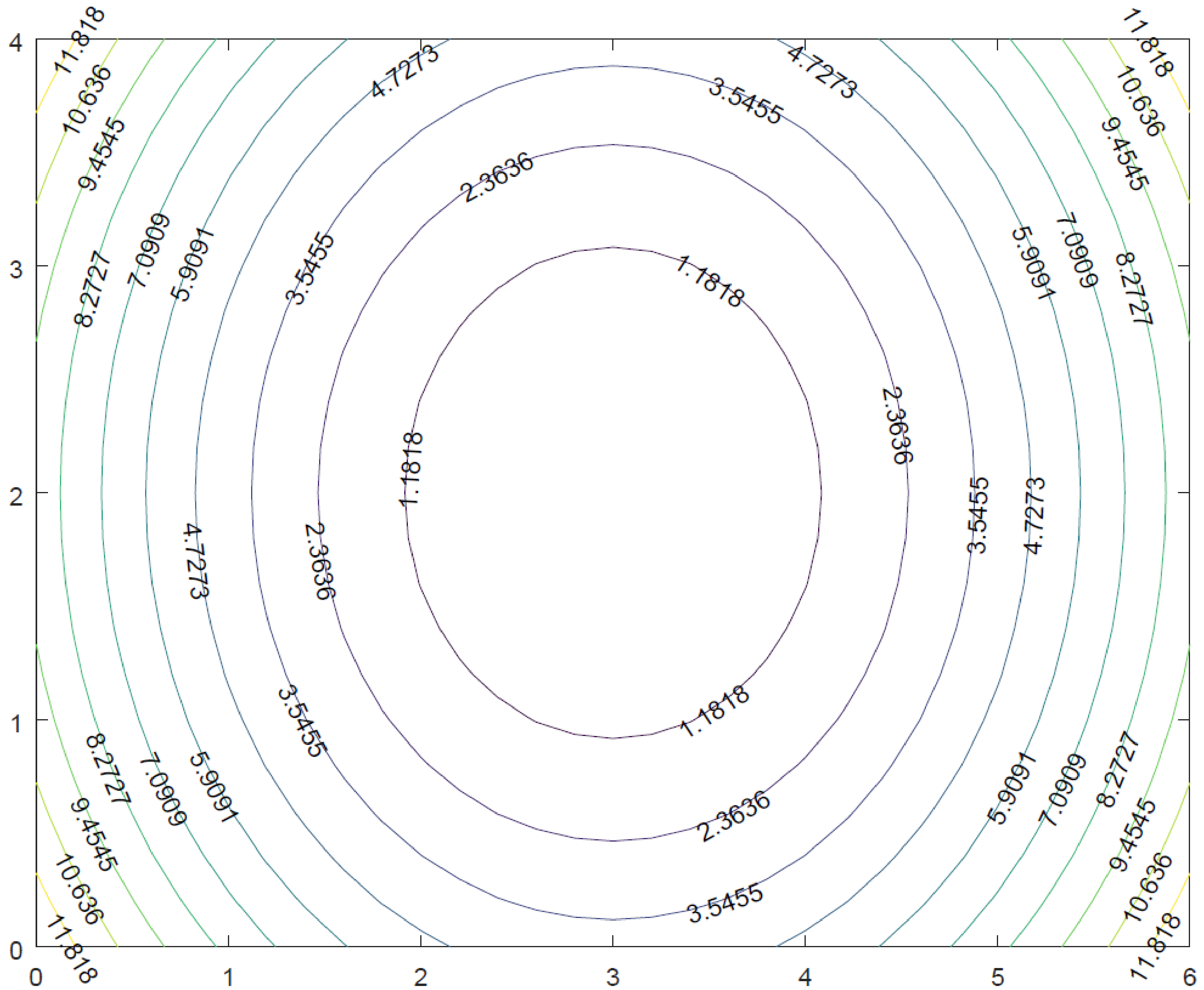
Surface



Contour Plot

```
home
x = 0:0.2:6;
y = 0:0.2:4;
[X,Y] = meshgrid(x,y);
Z = (X.-3).^2+(Y.-2).^2;
figure
contour(X,Y,Z,'ShowText','on')
```

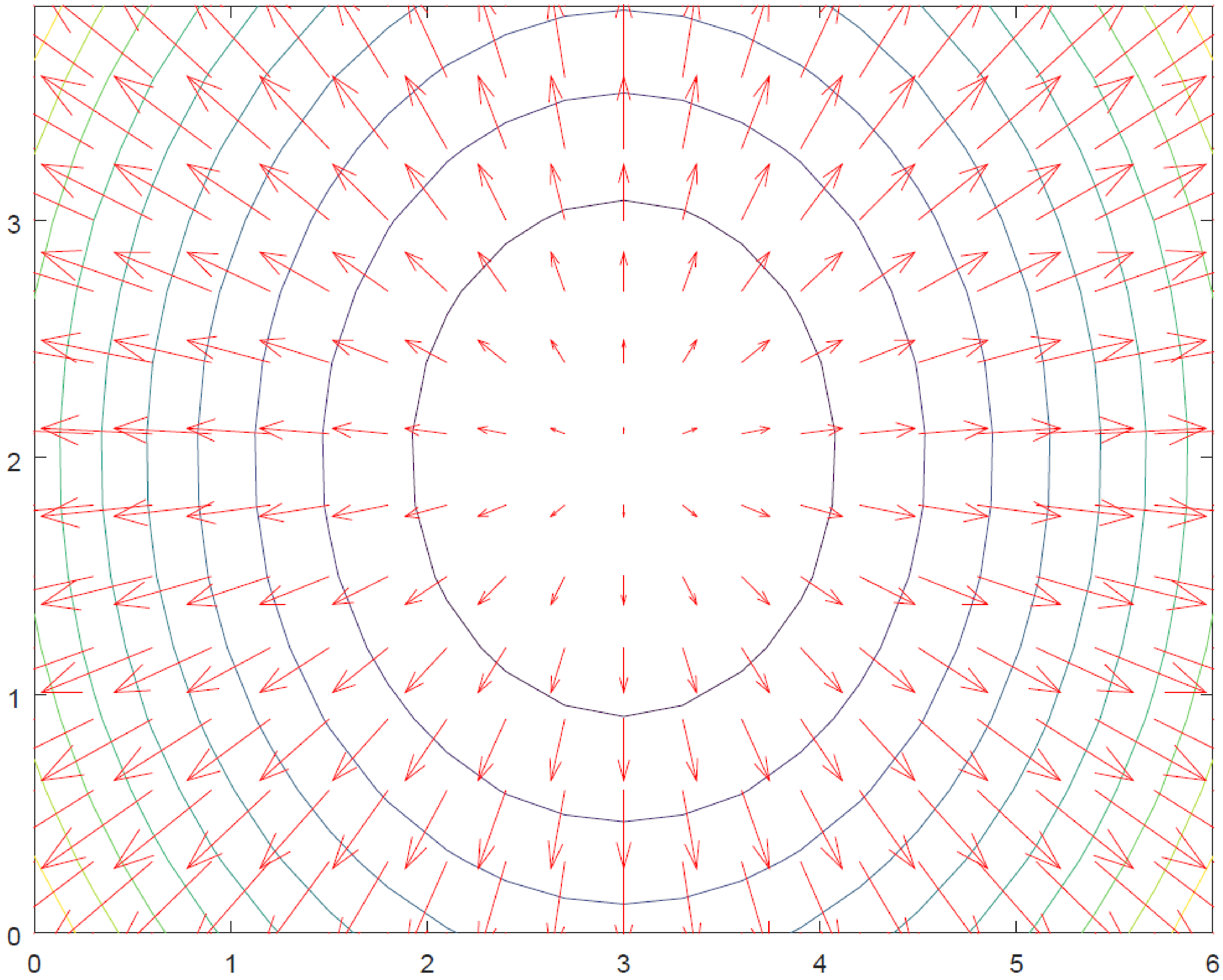
Contour Plot



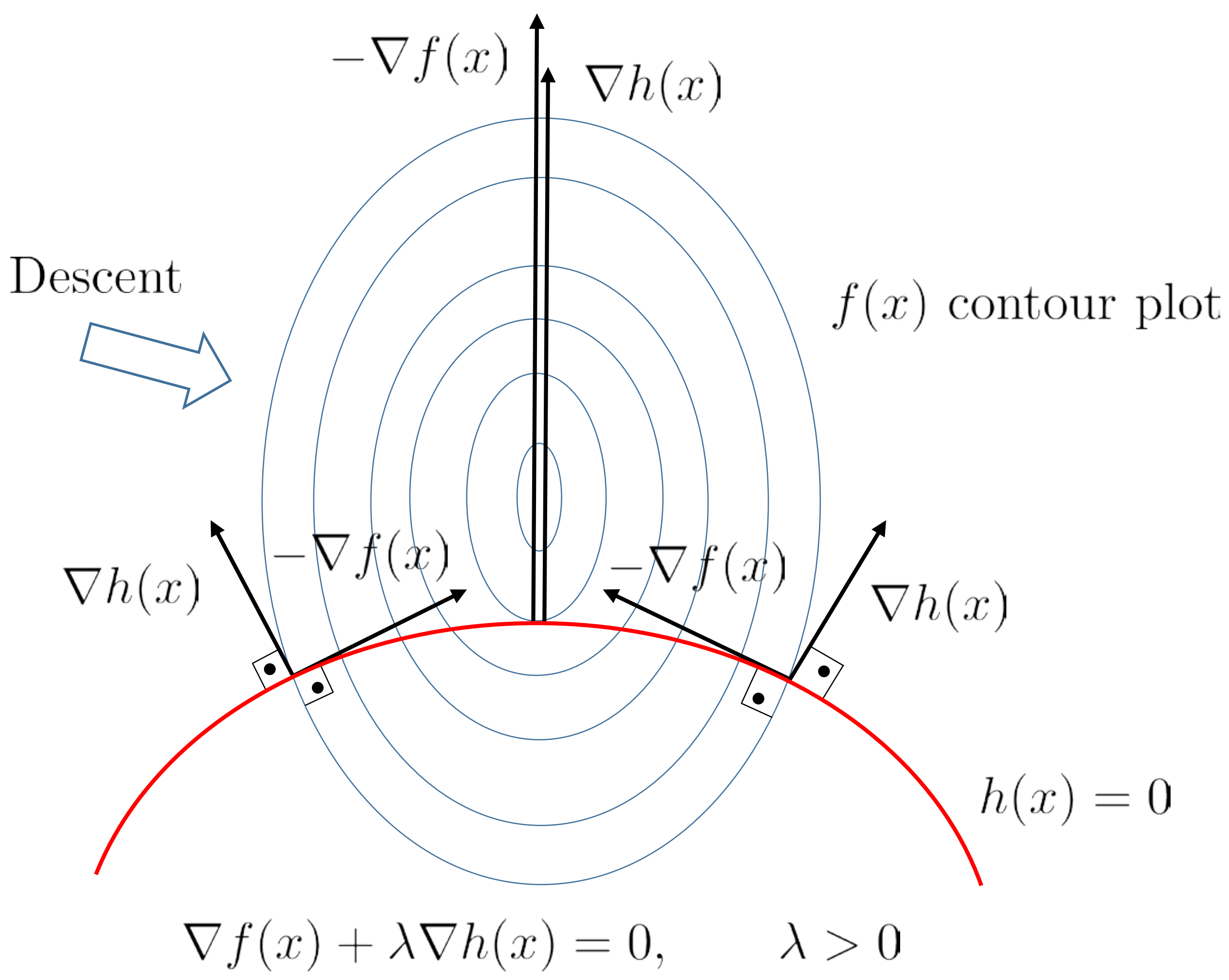
Contour Plot

```
x = 0:0.3:6;
y = 0:0.3:4;
[X,Y] = meshgrid(x,y);
Z = (X.-3).^2+(Y.-2).^2;
figure
contour(X,Y,Z,'ShowText','off')
[U,V] = gradient(Z);
hold on
quiver(X,Y,U,V,2,'color','red')
hold off
```

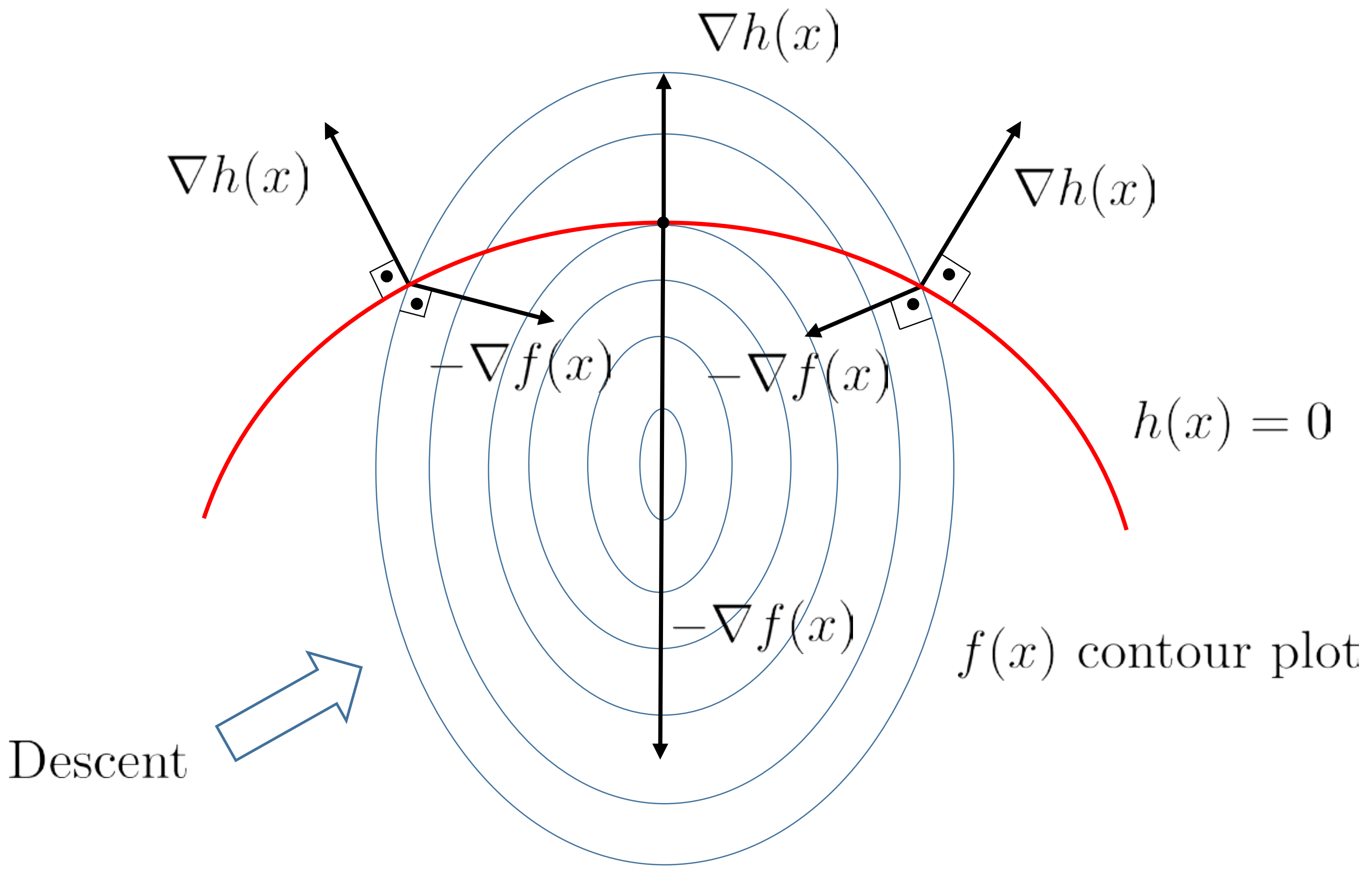
Contour Plot



FONC: Geometric Interpretation 1

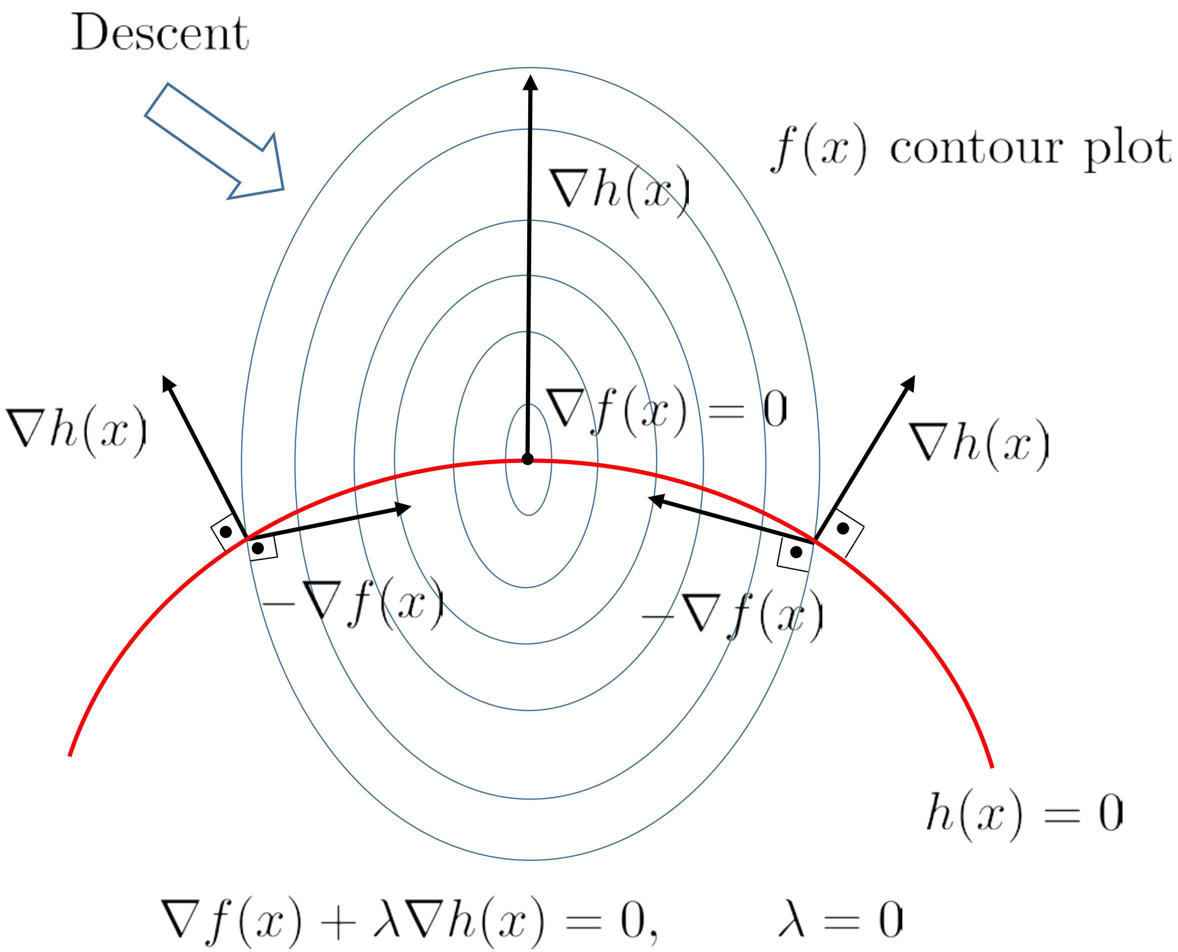

$-\nabla f(x)$
$\nabla h(x)$
Descent
$f(x)$ contour plot
$\nabla h(x)$
$-\nabla f(x)$
$-\nabla f(x)$
$\nabla h(x)$
$h(x) = 0$
$\nabla f(x) + \lambda \nabla h(x) = 0, \qquad \lambda > 0$

FONC: Geometric Interpretation 2



$$\nabla f(x) + \lambda \nabla h(x) = 0, \qquad \lambda < 0$$

FONC: Geometric Interpretation 3



Descent
$f(x)$ contour plot
$\nabla h(x)$
$\nabla f(x) = 0$
$\nabla h(x)$
$\nabla h(x)$
$-\nabla f(x)$
$-\nabla f(x)$
$h(x) = 0$
$\nabla f(x) + \lambda \nabla h(x) = 0, \qquad \lambda = 0$

Regularity

Regularity

The FONC for equality-constrained problems has one additional technical requirement, which is known as regularity.

A point is said to be **regular** if the gradients of the constraint functions at that point are all linearly independent.

Regularity

As the following example demonstrates, problems can have local minima that do not satisfy the regularity requirement, in which case they may violate the FONC.

Regularity: Example

Consider the equality-constrained problem:

$$\begin{aligned}
\min_{x} \ & f(x) = 2x_1 + 2x_2 \\
\text{s.t.} \ & h_1(x) = (x_1 - 1)^2 + x_2^2 - 1 = 0 \\
& h_2(x) = (x_1 + 1)^2 + x_2^2 - 1 = 0.
\end{aligned}$$

Regularity: Example

To apply the FONC to this problem, we define two Lagrange multipliers, λ_1 and λ_2, associated with the two constraints. The FONC and constraints of the problem are:

$$2 + 2\lambda_1(x_1 - 1) + 2\lambda_2(x_1 + 1) = 0$$

$$2 + 2\lambda_1 x_2 + 2\lambda_2 x_2 = 0$$

$$(x_1 - 1)^2 + x_2^2 - 1 = 0$$

$$(x_1 + 1)^2 + x_2^2 - 1 = 0.$$

Regularity: Example

Simultaneously solving the two constraints gives $(x_1, x_2) = (0, 0)$. Substituting these values into the FONC gives:

$$2 - 2\lambda_1 + 2\lambda_2 = 0$$

$$2 = 0,$$

which clearly has no solution.

Regularity: Example

However, $(x_1, x_2) = (0, 0)$ is the only feasible solution in the constraints, thus it is by definition a global and local minimum. This means that this problem has a local minimum that does not satisfy the FONC.

SOSC

SOSC

We state below sufficiency conditions.

Solving these conditions is equivalent to solving the problem under consideration.

SOSC

A solution x of the general problem meeting FONC conditions, where $f(x)$ is a continuously differentiable convex function and $h(x) = 0$ constitutes a convex set, is guaranteed to be an optimal solution.

SOSC

Additionally, the sufficiency condition below guarantees x to be an optimal solution.

Sufficiency conditions

$$\nabla_x^2 f(x) + \lambda^T \nabla_x^2 h(x) > 0$$

on the subspace

$$\{y : \nabla_x h(x) y = 0\}$$

Sufficiency conditions

Note that

$$\{y : \nabla_x h(x) y = 0\}$$

is the tangent hyperplane at x, and that

$$\lambda^T \nabla_x^2 h(x) = \sum_{i=1}^{m_E} \lambda_i^T \nabla_x^2 h_i(x)$$

Sufficiency conditions

In words, the condition above implies that the Hessian of the Lagrangian at x should be positive definite on the tangent hyperplane at x.

Observe that sufficiency conditions are second order conditions as they rely on second derivative matrices (Hessians).

This is it!