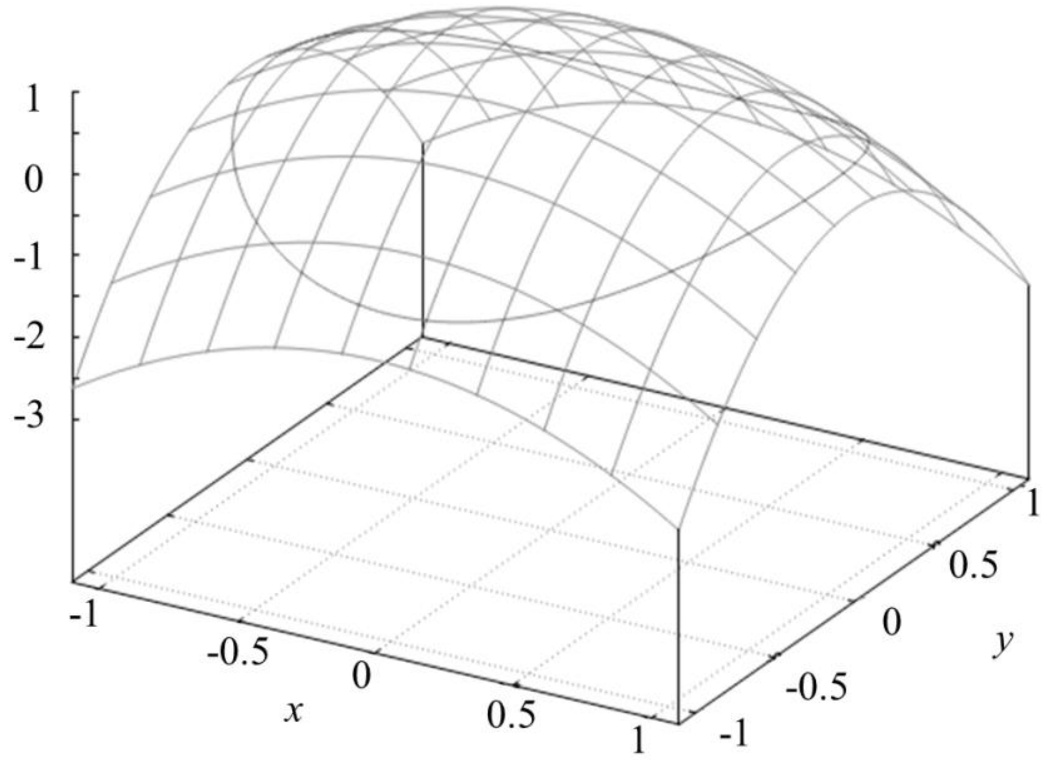
The Lagrange method of undetermined multipliers

Determining the extremum (maximum or minimum) of a multi-variable function:

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



Simultaneously find the extremum with respect to each separate variable

$$\begin{cases} \dfrac{df(x,y)}{dx} = -2x = 0 \\ \dfrac{df(x,y)}{dy} = -4y = 0 \end{cases}$$

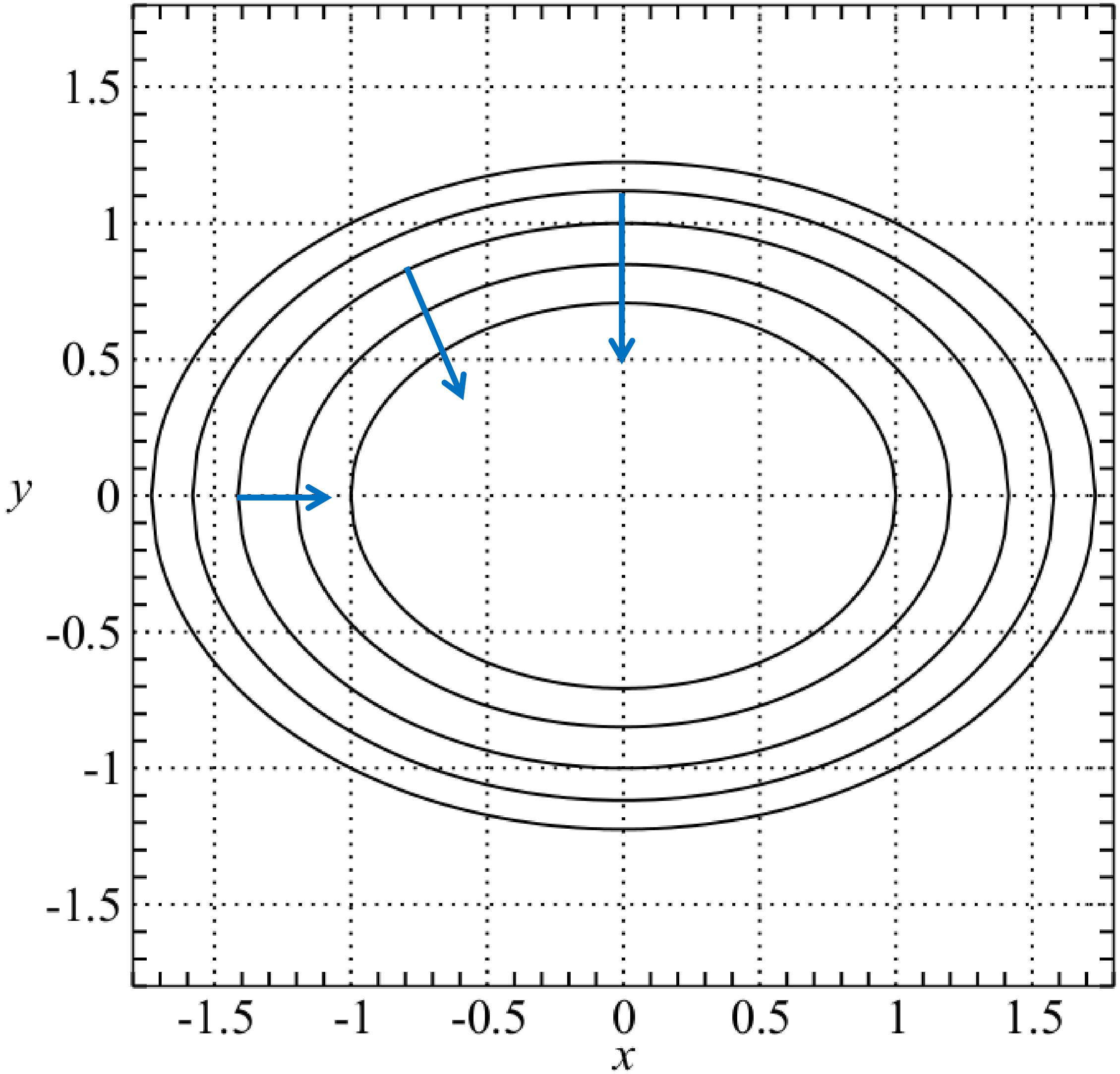
which gives the origin (0,0) as the maximum of this specific function (critical points and saddle points may arise in other cases)

Formally, this is equivalent to finding the point where the gradient vector is zero

$$\vec{\nabla} f(x,y) = \hat{\mathbf{i}}\frac{df(x,y)}{dx} + \hat{\mathbf{j}}\frac{df(x,y)}{dy} = 0$$

Contours of the surface projected onto the *xy*-plane

The contours represent curves of constant $f(x,y)$ projected onto the xy-plane



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

The gradient vectors show the direction of the greatest slope at each point

$$\vec{\nabla} f(x,y) = -2x\hat{\mathbf{i}} - 4y\hat{\mathbf{j}}$$

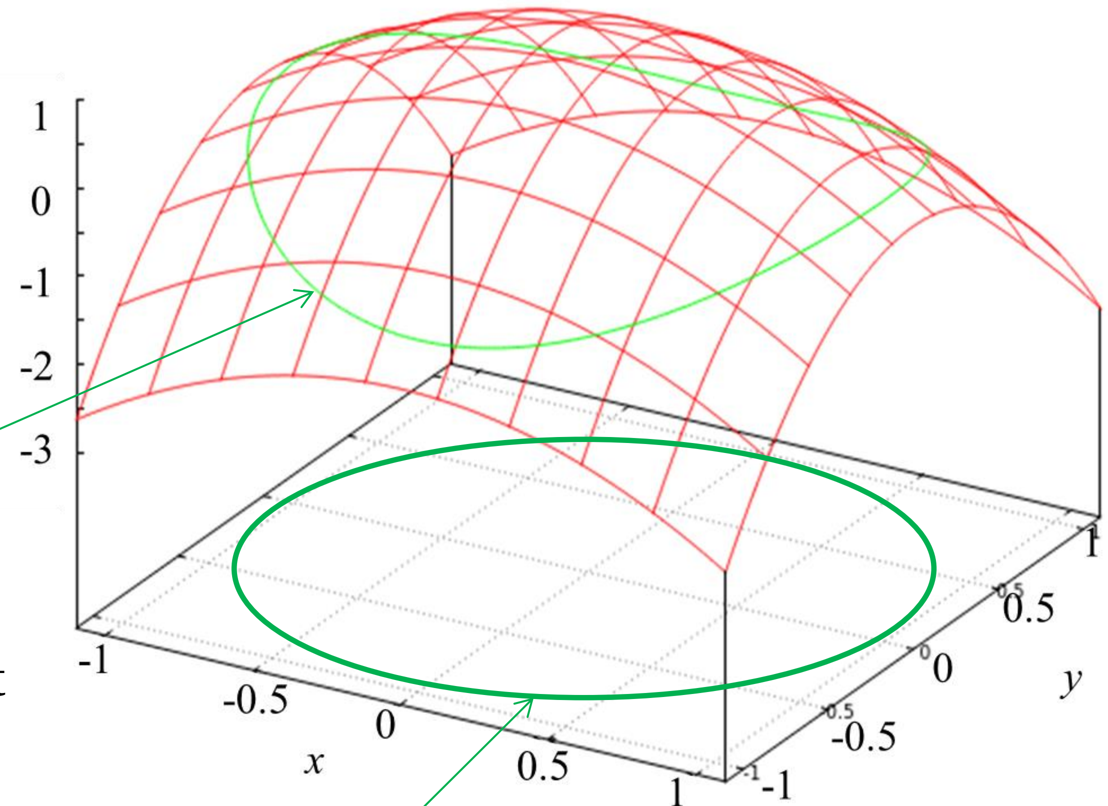
The gradient vector is 0 at the maximum or minimum of the function

Finding the maximum of a function subject to a constraint

How do we find the maximum of a function that simultaneously satisfies a constraint $g(x,y) = 0$, on the values of x and y?

Example: $g(x,y) = x^2 + y^2 - 1 = 0$

The optimization procedure has to be modified.



Projection of constraint onto the surface $f(x,y)$

Constraint on values of x and y

First method:

1) solve for y and x in the constraint equation
2) substitute in $f(x,y)$
3) optimize for remaining variable

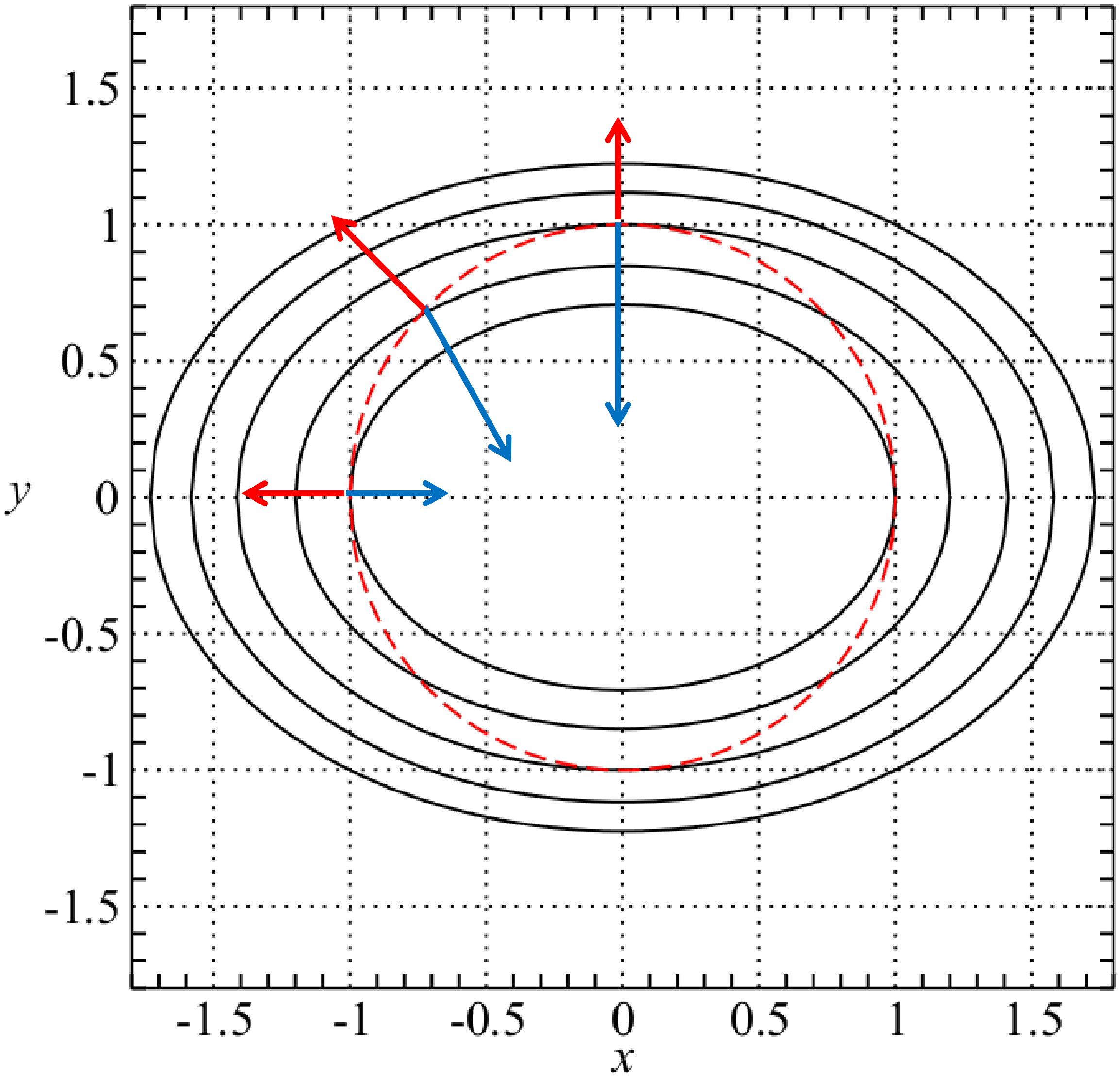
$$y = \pm\sqrt{1-x^2} \Longrightarrow f(x) = 2 - x^2 - 2\left(1-x^2\right) = x^2 \qquad (0,\pm 1)$$

$$x = \pm\sqrt{1-y^2} \Longrightarrow f(y) = 2 - (1-y^2) - 2y^2 = 1 - y^2 \qquad (\pm 1, 0)$$

Method of Lagrange Multipliers

Gradient at a contour line

Gradient of the constraint



In general, at points of intersection of the surface and constraint curve, the directions of the gradients are not the same.

If the gradient at the surface is parallel to the gradient of the constraint, we have a maximum / minimum point!

$$\vec{\nabla} f(x,y) = \lambda \vec{\nabla} g(x,y)$$

A scalar (value is related to the constraint)

$$\vec{\nabla} g(x,y) = 2x\hat{\mathbf{i}} + 2y\hat{\mathbf{j}} \qquad \vec{\nabla} f(x,y) = -2x\hat{\mathbf{i}} - 4y\hat{\mathbf{j}}$$

Method of Lagrange Multipliers

Define a new function $\Lambda(x,y,\lambda)$ of three variables for the constrained optimization:

$$\Lambda(x,y,\lambda) = f(x,y) + \lambda g(x,y) = 2 - x^2 - 2y^2 + \lambda(x^2 + y^2 - 1)$$

All equations at the constraint are satisfied when:

$$\vec{\nabla}\Lambda(x,y,\lambda) = 0$$

This gives three equations for the variables:

$$\begin{cases} \dfrac{d\Lambda(x,y,\lambda)}{dx} = -2x + 2\lambda x = 0 \\ \dfrac{d\Lambda(x,y,\lambda)}{dy} = -4y + 2\lambda y = 0 \\ \dfrac{d\Lambda(x,y,\lambda)}{d\lambda} = x^2 + y^2 - 1 = 0 \end{cases}$$

Solutions:

$(\lambda,x,y) = (-1,\pm1,0)$ and $(+2,0,\pm1)$,

Check that the gradients are parallel at the solutions:

$$\vec{\nabla}f(x,y) = -2x\hat{\mathbf{i}} - 4y\hat{\mathbf{j}} \qquad \vec{\nabla}g(x,y) = 2x\hat{\mathbf{i}} + 2y\hat{\mathbf{j}}$$

Method of Lagrange Multipliers

In a general case of a function $f(x_1, x_2, \ldots, x_N)$ subject to the α constraints $g_\nu(x_1, x_2, \ldots, x_N) = 0$ for $\nu = 1, 2, \ldots, \alpha$,
the procedure is analogous.

Define a new function $\Lambda(x_1, \ldots, x_N, \lambda_1, \ldots, \lambda_\alpha)$ of the $N+\alpha$ variables:

$$\Lambda(x_1, x_2, \cdots, x_N, \lambda_1, \lambda_2, \cdots, \lambda_\alpha) = f(x_1, x_2, \cdots, x_N) + \sum_{\nu=1}^{\alpha} \lambda_\nu g_\nu(x_1, x_2, \cdots, x_N)$$

All equations at the constraint are satisfied when:

$$\vec{\nabla}\Lambda(x_1, x_2, \cdots, x_N, \lambda_1, \lambda_2, \cdots, \lambda_\alpha) = 0 \Longrightarrow \begin{cases} \dfrac{d\Lambda(x_1, \cdots, x_N, \lambda_1, \cdots, \lambda_\alpha)}{dx_1} = 0 \\ \vdots \\ \dfrac{d\Lambda(x_1, \cdots, x_N, \lambda_1, \cdots, \lambda_\alpha)}{dx_N} = 0 \\ \dfrac{d\Lambda(x_1, \cdots, x_N, \lambda_1, \cdots, \lambda_\alpha)}{d\lambda_1} = 0 \\ \vdots \\ \dfrac{d\Lambda(x_1, \cdots, x_N, \lambda_1, \cdots, \lambda_\alpha)}{d\lambda_\alpha} = 0 \end{cases}$$

Solve the equations simultaneously to get the $x_1, x_2, \ldots, x_N$ and $\lambda_1, \lambda_2, \ldots, \lambda_\alpha$ variables

Excellent tutorial on Lagrange Multipliers

Khan Academy: **Lagrange multipliers, introduction**