Optimization with side constraints

Philipp Warode

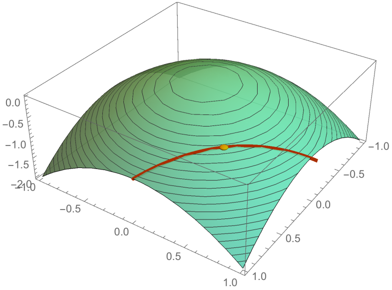
October 2, 2019

We want to maximize

$$\max -x^2 - y^2$$

with the side constraint

$$x + y = 1$$



- Reduce numbers of variables by inserting side constraints in the objective function

Example:
Side constraint:

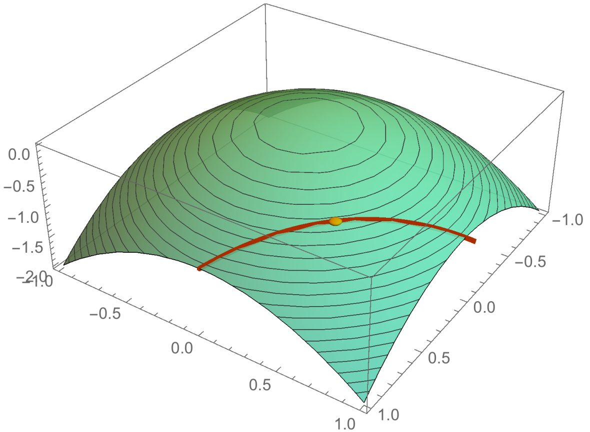
$$x + y = 1 \Leftrightarrow y = 1 - x$$

Objective function:

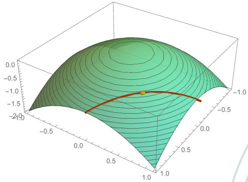
$$\max -x^2 - y^2 = \max -x^2 - (1 - x)^2$$

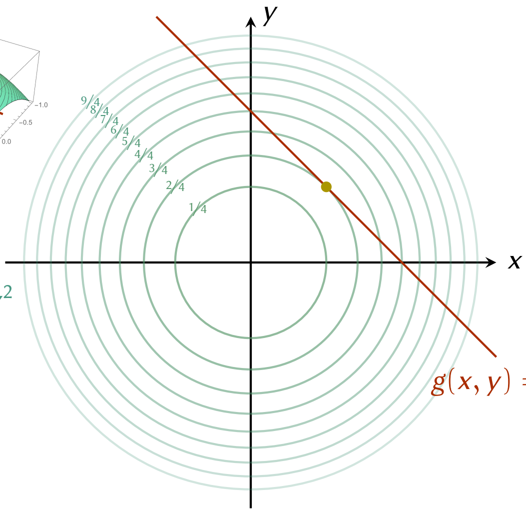
- Not always possible, e.g. if the side constraint can't be solved uniquely:

$$x^2 + y^2 = 1$$


0.0
−0.5
−1.0
−1.5
−2.0
−1.0
−0.5
0.0
0.5
1.0
−1.0
−0.5
0.0
0.5
1.0

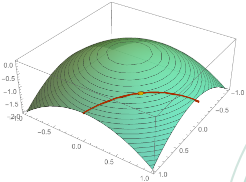
Lagrange Multipliers

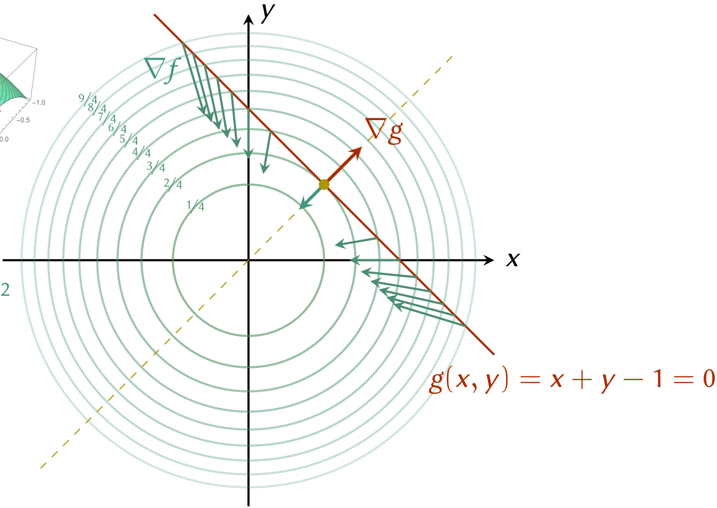




$$f(x,y) = -x^2 - y^2$$

$$g(x,y) = x + y - 1 = 0$$

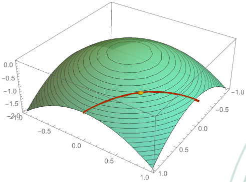


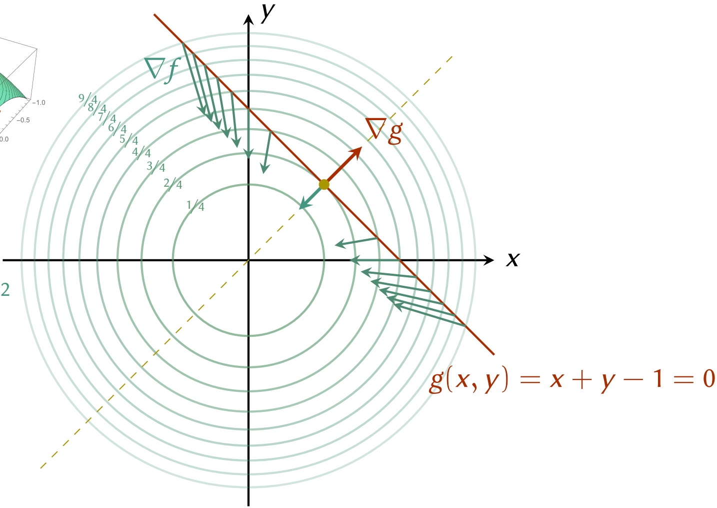


$f(x, y) = -x^2 - y^2$

- In the maximum, ∇f (gradient of the objective) and ∇g (gradient of the constraint) are linear dependent

Lagrange Multipliers





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

- In the maximum, ∇f (gradient of the objective) and ∇g (gradient of the constraint) are linear dependent
- Necessary condition for local optimum: $\nabla f = \lambda \nabla g$

Theorem

Let $x^{(0)} = (x_n^{(0)}, \dots, x_n^{(0)})$ *be an optimal solution of*

$$\begin{aligned} \max \quad & f(x) \\ \textit{s.t.} \quad & g_j(x) = 0 \quad \forall j \in \{1, \dots, m\} \end{aligned}$$

and f, g *partially differentiable with* $\nabla g_j(x^{(0)}) \neq 0$ *for* $j = 1, \dots, m$ *then there are* $\lambda_1, \dots, \lambda_m$ *such that*

$$\nabla f(x^{(0)}) + \sum_{j=1}^{m} \lambda_j \nabla g_j(x^{(0)}) = 0.$$

1. Create Lagrange function

$$\mathcal{L}(x,\lambda) := f(x) + \sum_{j=1}^{m} \lambda_j g_j(x)$$

1. Create Lagrange function

$$\mathcal{L}(x,\lambda) := f(x) + \sum_{j=1}^{m} \lambda_j g_j(x)$$

2. Compute the partial derivatives

$$\frac{\partial}{\partial x_i}\mathcal{L}(x,\lambda) = \frac{\partial}{\partial x_i} f(x) + \sum_{j=1}^{m} \lambda_j \frac{\partial}{\partial x_i} g_j(x)$$

$$\frac{\partial}{\partial \lambda_j}\mathcal{L}(x,\lambda) = g_j(x)$$

1. Create Lagrange function

$$\mathcal{L}(x,\lambda) := f(x) + \sum_{j=1}^{m} \lambda_j g_j(x)$$

2. Compute the partial derivatives

$$\frac{\partial}{\partial x_i}\mathcal{L}(x,\lambda) = \frac{\partial}{\partial x_i} f(x) + \sum_{j=1}^{m} \lambda_j \frac{\partial}{\partial x_i} g_j(x)$$

$$\frac{\partial}{\partial \lambda_j}\mathcal{L}(x,\lambda) = g_j(x)$$

3. Solve the system of equations

$$\frac{\partial}{\partial x_i}\mathcal{L}(x,\lambda) = 0 \text{ for all } i \text{ and } \frac{\partial}{\partial \lambda_j}\mathcal{L}(x,\lambda) = 0 \text{ for all } j$$