MATH529 – Fundamentals of Optimization

Constrained Optimization I

Marco A. Montes de Oca

Mathematical Sciences, University of Delaware, USA

Motivating Example

Motivating Example

min $cost(\mathbf{b}) = 1000 + 31b_1 + 20b_2 + 50b_3 + \ldots + 40b_n$
(EUR)

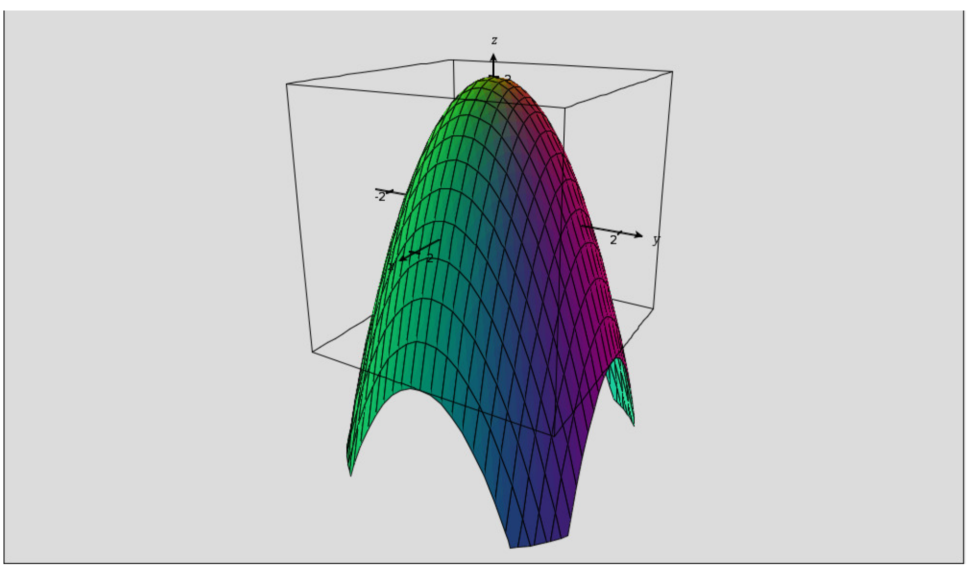
subject to:

$b_1 + b_2 + \ldots + b_n \leq 6500$ (Casks)

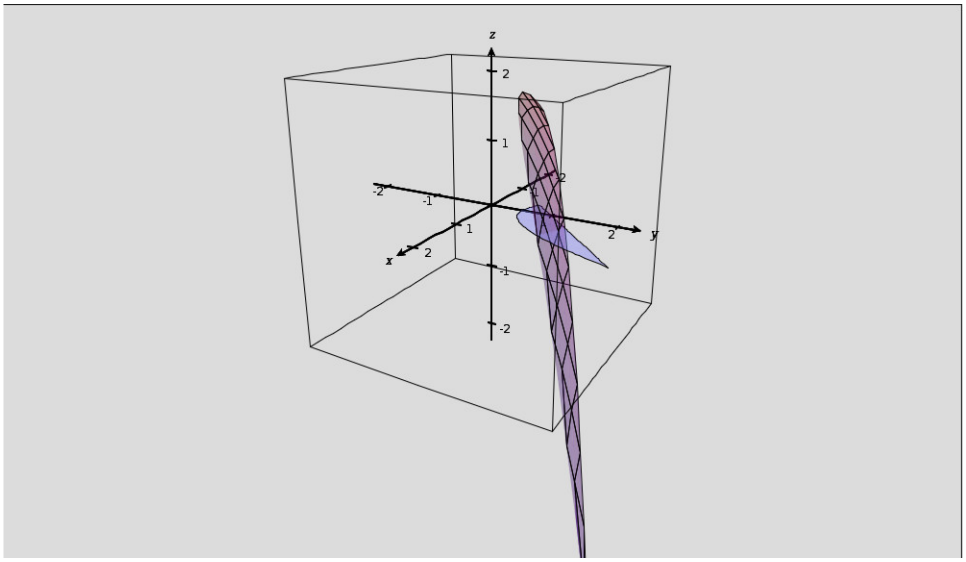
$0.6b_1 + 0.5b_2 \leq 30$ (Tons of ingredient 1)
$0.3b_1 + 0.1b_2 + 0.4b_4 + 0.5b_{12} \leq 20$ (Tons of ingredient 2)
$\vdots$
$0.1b_4 + 0.05b_6 \leq 5$ (Tons of ingredient n)

$b_i \geq 100$, $i \in \{1, 2, \ldots, n\}$ (Casks) (minimum production).

Effects of constraints



Effects of constraints



Definitions

A general model of a constrained optimization problem is:

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

subject to

$$c_i(x) = 0, \quad i \in \mathcal{E}$$

$$c_i(x) \geq 0 \text{ (or } c_i(x) \leq 0), \quad i \in \mathcal{I}$$

where f is called the *objective function*, the functions $c_i(x), \ i \in \mathcal{E}$ are the *equality constraints*, and the functions $c_i(x), \ i \in \mathcal{I}$ are the *inequality constraints*.

Definitions

The feasible set $\Omega \subset \mathbb{R}^n$ is the set of points that satisfy the constraints:

$$\Omega = \{x \mid c_i(x) \geq 0, \;\; i \in \mathcal{I}, c_i(x) = 0, \;\; i \in \mathcal{E}\}$$

Therefore, a constrained optimization problem can be defined simply as

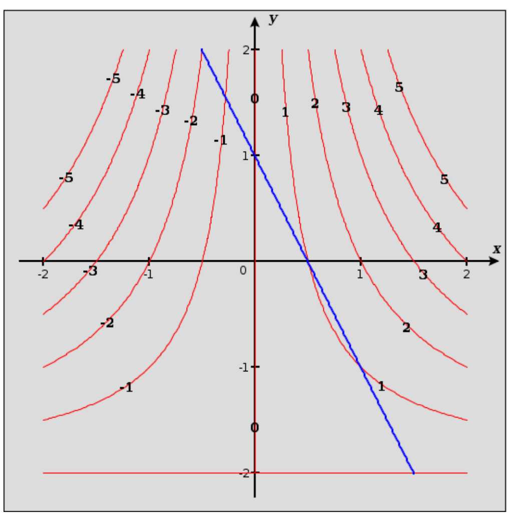
$$\min_{x \in \Omega} f(x)$$

Example: One equality constraint

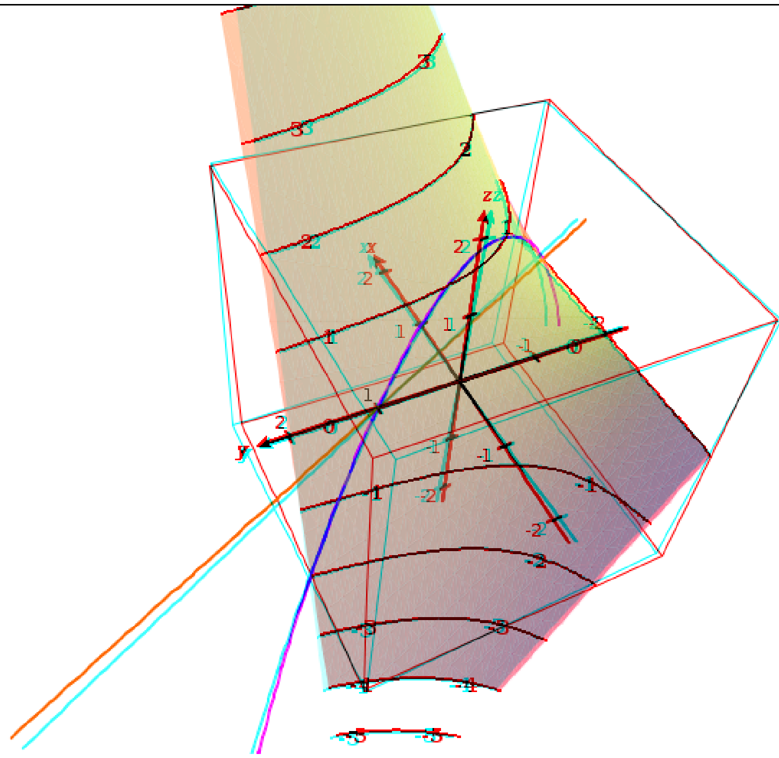
Suppose you want to solve

$\max f(x) = x_1 x_2 + 2x_1,$

subject to $g(x) = 2x_1 + x_2 = 1.$



Example: One equality constraint



Example: One equality constraint

Suppose you want to solve

$\max f(x) = x_1x_2 + 2x_1$,

subject to $g(x) = 2x_1 + x_2 = 1$.

Solution method 1: Substitution

From the constraint, we see that $x_2 = 1 - 2x_1$. Thus, $f(x)$ can be rewritten as $f(x_1) = x_1(1 - 2x_1) + 2x_1$.

Example: One equality constraint

Suppose you want to solve

$\max f(x) = x_1 x_2 + 2x_1$,

subject to $g(x) = 2x_1 + x_2 = 1$.

Solution method 1: Substitution

From the constraint, we see that $x_2 = 1 - 2x_1$. Thus, $f(x)$ can be rewritten as $f(x_1) = x_1(1 - 2x_1) + 2x_1$.

This method works only on very few cases. E.g., it does not work when one cannot solve for one of the variables.

Example: One equality constraint

Solution method 2: Lagrange multipliers

Define the *Lagrangian function*:

$Z = f(x) + \lambda(0 - c_1(x)) = x_1 x_2 + 2x_1 + \lambda(1 - 2x_1 - x_2)$

where λ is a so-called *Lagrange multiplier*.

If the constraint is satisfied, then $c_1(x) = 0$, and Z is identical to f.

Therefore, as long as $c_1(x) = 0$, searching for the maximum of Z is the same as searching for the maximum of f.

So, **how to always satisfy** $c_1(x) = 0$**?**

Example: One equality constraint

Solution method 2: Lagrange multipliers

If $Z = Z(\lambda, x_1, x_2)$, then $\nabla Z = 0$ implies

$\frac{\partial Z}{\partial \lambda} = 1 - 2x_1 - x_2 = 0$, or simply $c_1(x) = 0$ (the original constraint)

$\frac{\partial Z}{\partial x_1} = x_2 + 2 - 2\lambda = 0$, and

$\frac{\partial Z}{\partial x_2} = x_1 - \lambda = 0$

Solving this system: $x_1 = 3/4$, $x_2 = -1/2$, and $\lambda = 3/4$. A second order condition should be used to tell whether $(3/4, -1/2)$ is a maximum or a minimum.

Lagrange multipliers

In general, to find an extremum $x^\star$ of $f(x)$ subject to $c_1(x) = 0$, we define the Lagrangian function:

$$L(x, \lambda) = f(x) - \lambda c_1(x)$$

Then, $\nabla L(x^\star, \lambda^\star) = 0$ implies

$c_1(x^\star) = 0$ and $\nabla f(x^\star) - \lambda^\star \nabla c_1(x^\star) = 0$, or equivalently $\nabla f(x^\star) = \lambda^\star \nabla c_1(x^\star)$.

Exercise

Find all the extrema of $f(x) = x_1^2 + x_2^2$, subject to $x_1^2 + x_2 = 1$

Lagrange Multiplier Method: Derivation

Total differential approach:

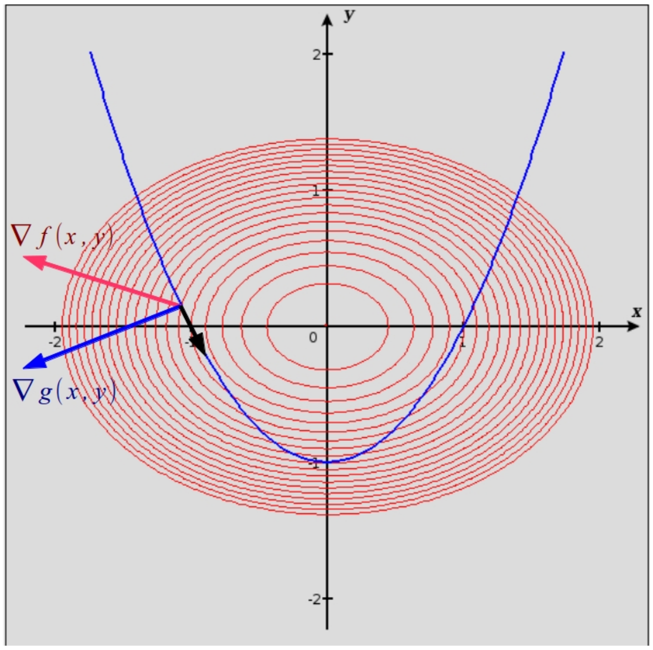
Let $z = f(x, y)$ (objective function) and $g(x, y) = c$ (equality constraint). At an extremum, the first order condition translates into $dz = 0$, so $dz = f_x dx + f_y dy = 0$ (1). Since dx and dy are not independent (due to the constraint), we can take the differential of g as well: $dg = g_x dx + g_y dy = 0$ (2). From (2), $dx = -\frac{g_y}{g_x} dy$. Substituting dx in (1):

$-f_x \frac{g_y}{g_x} dy + f_y dy = 0$, which implies $\frac{f_x}{g_x} = \frac{f_y}{g_y}$.

If $\frac{f_x}{g_x} = \frac{f_y}{g_y} = \lambda$, then $f_x - \lambda g_x = 0$ and $f_y - \lambda g_y = 0$, which are the Lagrange multiplier equations.

Lagrange Multiplier Method: Derivation

Taylor series approach:



Lagrange Multiplier Method: Derivation

Taylor series approach:

If we want to satisfy the constraint as we move from x to $x + s$, then

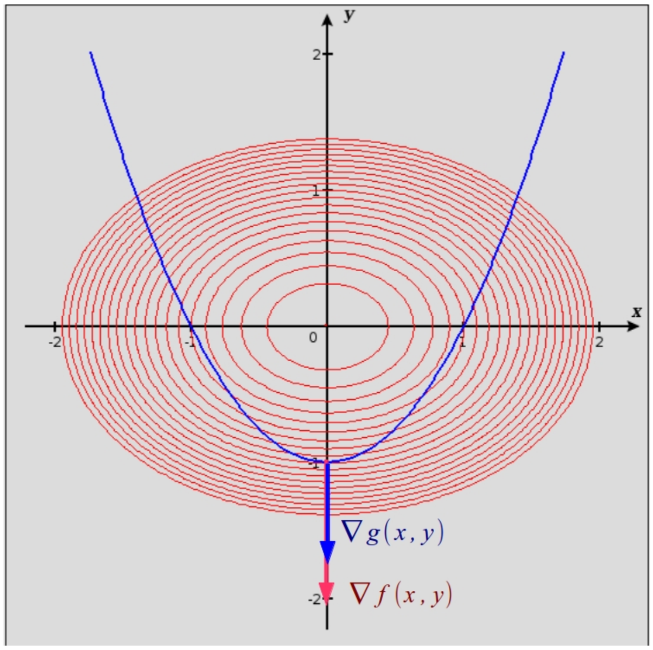
$0 = c_1(x + s) \approx c_1(x) + \nabla c_1(x)^T s$, but since $c_1(x) = 0$, then $\nabla c_1(x)^T s = 0$ (1)

Additionally, if we want to decrease the function as we move, we would require $\nabla f(x)^T s < 0$ (2).

Only when $\nabla f(x) = \lambda \nabla g(x)$, we **cannot** find s to satisfy (1) and (2).

Lagrange Multiplier Method: Derivation

Taylor series approach:



Shadow Price

The Lagrange multiplier measures the sensitivity of the optimal solution to changes in the constraint. For example, assume that $\lambda = \lambda(B)$, $x = x(B)$, and $y = y(B)$. If you want to maximize $U(x)$ subject to $g(x) = B$ (so that $c_1(x) = B - g(x) = 0$). Then,

$L(x, \lambda) = U(x) + \lambda(B - g(x))$. By the Chain Rule:

$$\frac{dL}{dB} = L_{x_1}\frac{dx_1}{dB} + L_{x_2}\frac{dx_2}{dB} + L_\lambda \frac{d\lambda}{dB}$$
$$\frac{dL}{dB} = (U_{x_1} - \lambda g_{x_1})\frac{dx_1}{dB} + (U_{x_2} - \lambda g_{x_2})\frac{dx_2}{dB} + (B - g(x))\frac{d\lambda}{dB} + \lambda(1).$$

Since the first order condition says that $U_{x_1} - \lambda g_{x_1} = 0$, $U_{x_2} - \lambda g_{x_2} = 0$, and $B - g(x) = 0$, we have

$\frac{dL}{dB} = \lambda$. (This equation answers the question "Will a slight relaxation of the budget constraint increase or decrease the optimal value of U?")

Example

Use the Lagrange-multiplier method to find stationary values of $z = x - 3y - xy$, subject to $x + y = 6$. Will a slight relaxation of the constraint increase or decrease the optimal value of z? At what rate?