Multivariate Calculus & Constrained Optimization

Lecture 14

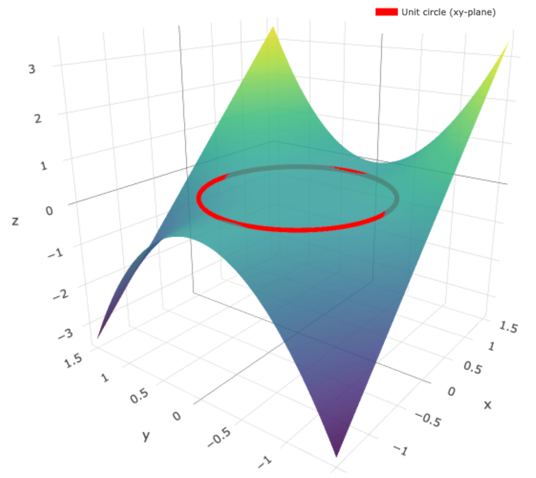
Termeh Shafie

1

the problem

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

on the set of all values $\underbrace{x^2 + y^2 = 1}_{\text{unit circle}}$



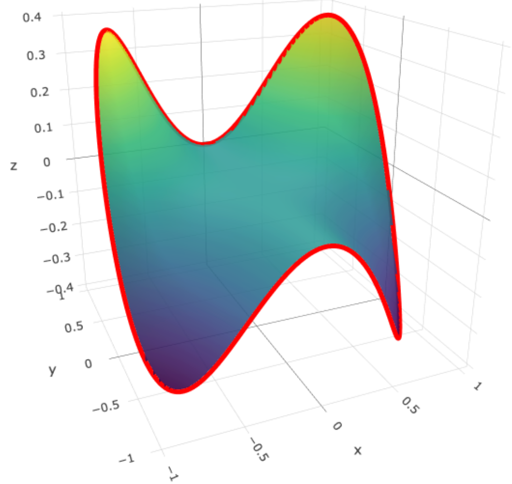
2

the problem

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

on the set of all values $\underbrace{x^2 + y^2 = 1}_{\text{unit circle}}$

when projected onto the surface



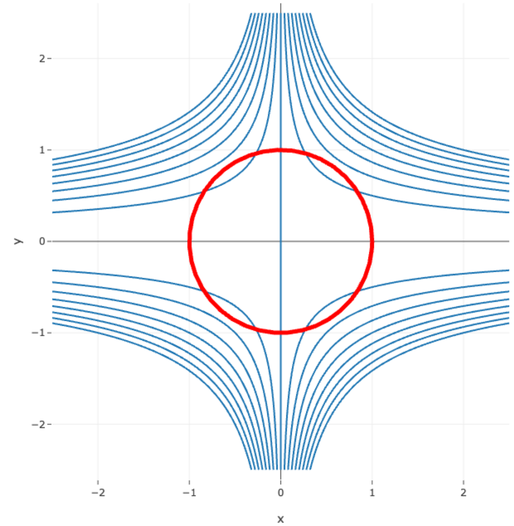
3

the problem

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

on the set of all values $\underbrace{x^2 + y^2 = 1}_{\text{unit circle}}$

easier to work with contour map
which limit the view to the input space



4

the problem

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

on the set of all values $\underbrace{x^2 + y^2 = 1}_{\text{unit circle}}$

easier to work with contour maps
which limit the view to the input space

contour level $c \iff \{(x, y) : f(x, y) = c\}$

the problem

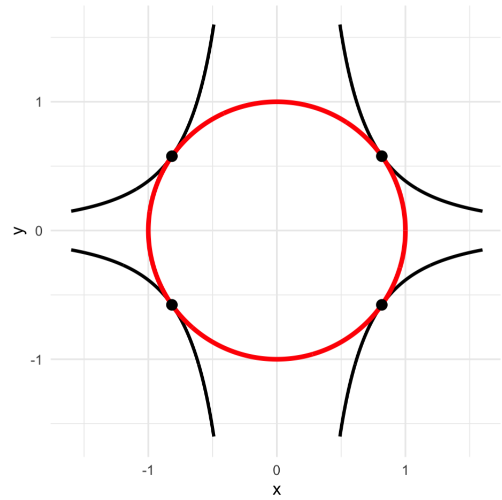
maximize $f(x, y) = x^2 y$

on the set of all values $\underbrace{x^2 + y^2 = 1}_{\text{unit circle}}$

easier to work with contour maps
which limit the view to the input space

contour level $c \iff \{(x, y) : f(x, y) = c\}$

`use the gradient`



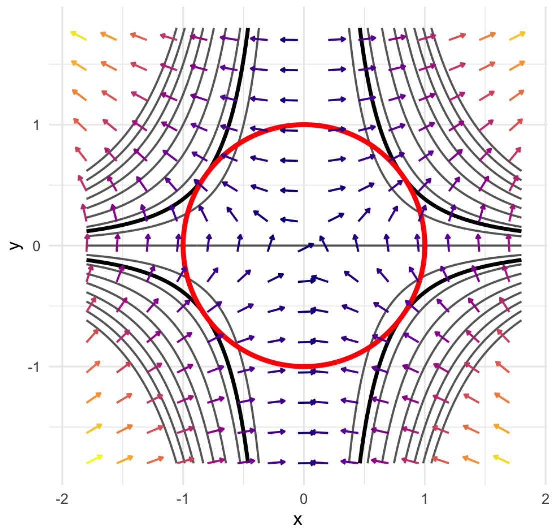
the problem

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

on the set of all values $\underbrace{x^2 + y^2 = 1}_{\text{unit circle}}$

$f(x, y) = c$

gradient vector are perpendicular
when crossing a contour line

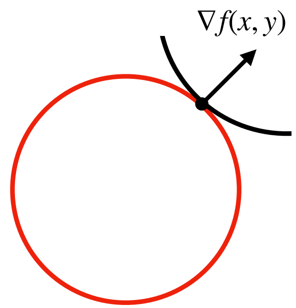


the problem

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

on the set of all values $\underbrace{x^2 + y^2 = 1}_{\text{unit circle}}$

$f(x, y) = c$



the problem

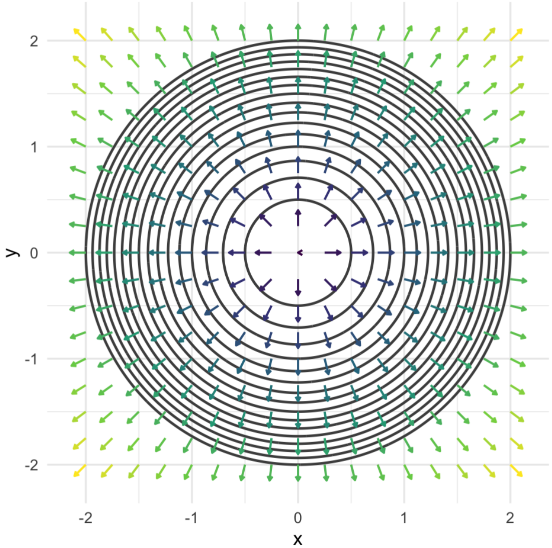
maximize $f(x, y) = x^2y$

on the set of all values $\underbrace{x^2 + y^2 = 1}_{\text{unit circle}}$

$f(x, y) = c$

now let $g(x, y) = x^2 + y^2$

gradient vector are perpendicular
when crossing a contour line



the problem

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

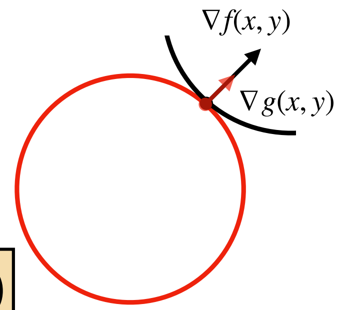
on the set of all values $\underbrace{x^2 + y^2 = 1}_{\text{unit circle}}$

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

∇g is proportional to ∇f

$$\nabla f\left(x_{\max}, y_{\max}\right) = \lambda \nabla g\left(x_{\max}, y_{\max}\right)$$

where λ is the Lagrange multiplier



solution to the problem

$$\nabla f\left(x_{\max}, y_{\max}\right) = \lambda \nabla g\left(x_{\max}, y_{\max}\right)$$

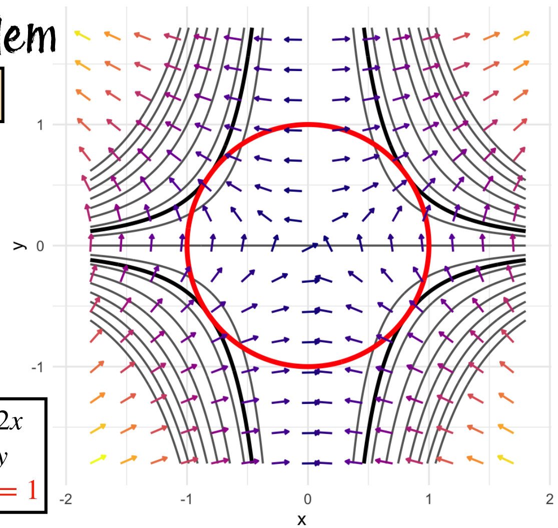
$f(x, y) = x^2y$

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) = \begin{bmatrix} 2xy \\ x^2 \end{bmatrix}$$

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

$$\nabla g = \left(\frac{\partial f}{\partial x}, \frac{\partial g}{\partial y}\right) = \begin{bmatrix} 2x \\ 2y \end{bmatrix}$$

$$\Longrightarrow \begin{bmatrix} 2xy \\ x^2 \end{bmatrix} = \lambda \begin{bmatrix} 2x \\ 2y \end{bmatrix} \Longrightarrow \begin{array}{l} 2xy = \lambda 2x \\ x^2 = \lambda 2y \\ x^2 + y^2 = 1 \end{array}$$



solution to the problem

$$\nabla f\left(x_{\max}, y_{\max}\right) = \lambda \nabla g\left(x_{\max}, y_{\max}\right)$$

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

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) = \begin{bmatrix} 2xy \\ x^2 \end{bmatrix}$$

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

$$\nabla g = \left(\frac{\partial f}{\partial x}, \frac{\partial g}{\partial y}\right) = \begin{bmatrix} 2x \\ 2y \end{bmatrix}$$

$$\Longrightarrow \begin{bmatrix} 2xy \\ x^2 \end{bmatrix} = \lambda \begin{bmatrix} 2x \\ 2y \end{bmatrix}$$

assuming $x \neq 0$:

$2xy = \lambda 2x \quad \Longrightarrow 2y = \lambda 2 \Longrightarrow y = \lambda$

$x^2 = \lambda 2y \quad \Longrightarrow x^2 = 2y^2$

$x^2 + y^2 = 1 \quad \Longrightarrow 2y^2 + y^2 = 1$

$3y^2 = 1$

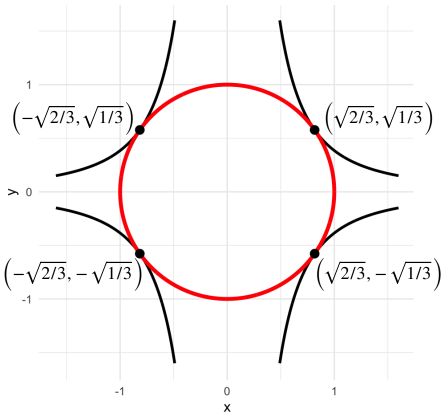
$y = \pm\sqrt{1/3}$

$\Longrightarrow x = 2y^2 = 2\left(\sqrt{1/3}\right)^2 \Longrightarrow x = \pm\sqrt{2/3}$

check: if $x = 0$ then $y = 0$, constraint is not satisfied

$\Longrightarrow$ *four possible points satisfying the constraint*

solution to the problem



assuming $x \neq 0$:

$$2xy = \lambda 2x \implies 2y = \lambda 2 \implies y = \lambda$$

$$x^2 = \lambda 2y \implies x^2 = 2y^2$$

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

$$3y^2 = 1$$

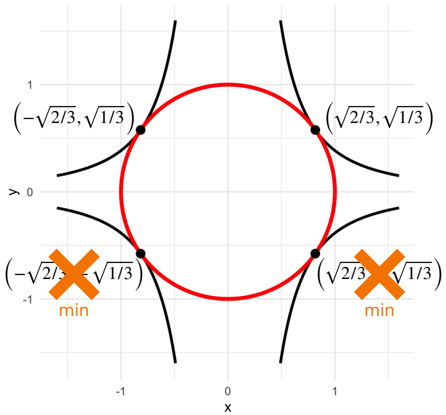
$$y = \pm\sqrt{1/3}$$

$$\implies x = 2y^2 = 2\left(\sqrt{1/3}\right)^2 \implies x = \pm\sqrt{2/3}$$

check: if $x = 0$ then $y = 0$, constraint is not satisfied

$\implies$ *four possible points satisfying the constraint*

solution to the problem



$$f(x, y) = \underbrace{x^2}_{\text{always positive}} y$$

$$f(-\sqrt{2/3}, \sqrt{1/3}) = \frac{2}{3}\sqrt{1/3}$$

$$f(\sqrt{2/3}, \sqrt{1/3}) = \frac{2}{3}\sqrt{1/3}$$

The Lagrangian

Let $f : \mathbb{R}^n \to \mathbb{R}$ be the objective function,

and $g_i : \mathbb{R}^n \to \mathbb{R}$ for $i = 1, \dots, m$ be constraint functions.

The Lagrangian is the function $\mathscr{L} : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ defined by

$$\mathscr{L}(x, \lambda) = f(x_1, \dots, x_n) + \sum_{i=1}^{m} \lambda_i g_i(x_1, \dots, x_n)\,.$$

where

$x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ are the primal variables,

$\lambda = (\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m$ are the Lagrange multipliers (dual variables).

Note: the primal variables are the original variables of the optimization problem
(e.g. while the Lagrange multipliers are auxiliary variables introduced to enforce the constraints.

Lagrangian: Equality Constraint

If the problem is

$$\text{maximize/minimize } f(x) \quad \text{subject to} \quad g(x) = 0,$$

the Lagrangian is

$$\mathscr{L}(x, \lambda) = f(x) + \lambda\, g(x)$$

for example: $x = (x_1, x_2) = (x, y) \implies \mathscr{L}(x, y, \lambda) = f(x, y) + \lambda g(x, y)\,.$

First-order (Lagrange multiplier) condition

Let $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, $\lambda = (\lambda_1, \ldots, \lambda_m) \in \mathbb{R}^m$.
A point (x^*, λ^*) is a stationary point of the Lagrangian if

$$\nabla_x \mathscr{L}(x^*, \lambda^*) = 0, \qquad \nabla_\lambda \mathscr{L}(x^*, \lambda^*) = 0,$$

which explicitly means

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla g_i(x^*) = 0, \qquad g_i(x^*) = 0 \quad \text{for } i = 1,\ldots,m$$

for example: $x = (x_1, x_2) = (x, y) \implies \nabla f(x^*, y^*) + \lambda^* \nabla g(x^*, y^*) = 0, \quad g(x^*, y^*) = 0.$

At an optimum under constraints, **the gradient of f lies in the span of the gradients of the constraints** i.e. level sets are **tangent**.

First-order (Lagrange multiplier) condition

Let $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, $\lambda = (\lambda_1, \ldots, \lambda_m) \in \mathbb{R}^m$.
A point (x^*, λ^*) is a stationary point of the Lagrangian if

$$\nabla_x \mathscr{L}(x^*, \lambda^*) = 0, \qquad \nabla_\lambda \mathscr{L}(x^*, \lambda^*) = 0,$$

which explicitly means

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla g_i(x^*) = 0, \qquad g_i(x^*) = 0 \quad \text{for } i = 1,\ldots,m$$

for example: $x = (x_1, x_2) = (x, y) \implies \nabla f(x^*, y^*) + \lambda^* \nabla g(x^*, y^*) = 0, \quad g(x^*, y^*) = 0.$

Note:
$\mathscr{L}(x, y, \lambda) = f(x, y) + \lambda\, g(x, y)$ with stationary points $\nabla f(x^*) + \lambda^* \nabla g(x^*) = 0$
is equivalent to
$\mathscr{L}(x, y, \lambda) = f(x, y) - \lambda\, g(x, y)$ with stationary points $\nabla f(x^*) - \lambda^* \nabla g(x^*) = 0.$

The Lagrangian

example

maximize $f(x, y) = x^2 y$
subject to the constraint $g(x, y) = x^2 + y^2 = 1$

Lagrangian: $\mathscr{L}(x, y, \lambda) = f(x, y) - \lambda[g(x, y) - 1] \implies \nabla \mathscr{L} = \begin{bmatrix} \frac{\partial \mathscr{L}}{\partial x} \\ \frac{\partial \mathscr{L}}{\partial y} \\ \frac{\partial \mathscr{L}}{\partial \lambda} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

$$\frac{\partial \mathscr{L}}{\partial x} = \frac{f(x, y) - \lambda[g(x, y) - 1]}{\partial x} = 0$$

$$\frac{\partial \mathscr{L}}{\partial y} = \frac{f(x, y) - \lambda[g(x, y) - 1]}{\partial y} = 0$$

generally: $g(x, y) = c$

$$\frac{\partial \mathscr{L}}{\partial \lambda} = \frac{f(x, y) - \lambda[g(x, y) - 1]}{\partial \lambda} = 0 - [g(x, y) - 1] = 0 \implies g(x, y) = x^2 + y^2 = 1$$

The Lagrangian

$\max_{x,y} f(x, y)$ subject to $g(x, y, c) = 0,$

where: c is now a parameter (not a choice variable) and x, y are the primal variables.

Define the Lagrangian $\mathscr{L}(x, y, \lambda; c) = f(x, y) + \lambda\, g(x, y, c)$.

At an optimum (x^*, y^*, λ^*), the first-order conditions are $\nabla_{x,y} \mathscr{L}(x^*, y^*, \lambda^*; c) = 0, \quad g(x^*, y^*, c) = 0.$

Define the optimal value: $M^* = f(x^*, y^*)$.

Since x^* and y^* depend on c we can write this explicitly as

$$M^*(c) = f\big(x^*(c), y^*(c)\big).$$

The envelope theorem states that $\lambda^* = \dfrac{dM^*}{dc} = \lambda^* \dfrac{\partial g(x^*, y^*, c)}{\partial c}$

The Lagrange multiplier equals the derivative of the value function with respect to the constraint parameter.

The Lagrangian

exercise

Suppose you run a small workshop. Labor costs €10 per hour and steel costs €40 per ton.

Revenue is modeled as $R(h, s) = 100\, h^{1/2} s^{1/2},$ where h is labor hours and s is tons of steel.

If your budget is €400, what is the maximum possible revenue? Is the constraint valuable to relax?