Computational Science:
Computational Methods in Engineering

Steepest Ascent Method

Outline

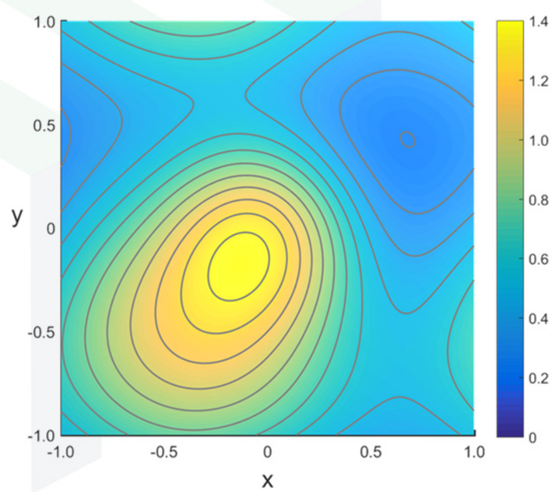
- Description of the Method
- Summary of the Method
- Choice of α
- Example

Description of the Method

Concept of Steepest Ascent

Suppose there is a function $f(x, y)$.

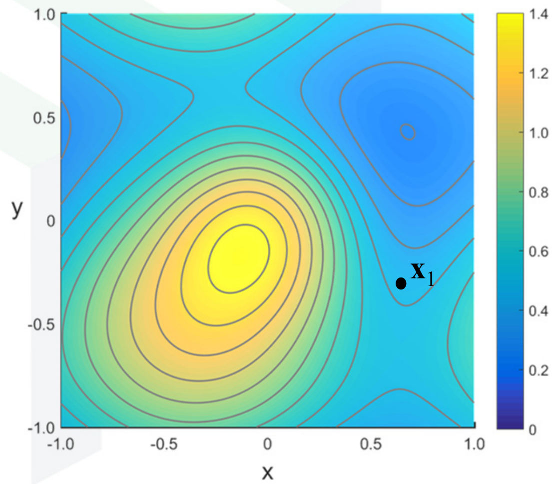
How is the maximum found?



EMPossible

Concept of Steepest Ascent

Step 1 – Pick a starting point $\mathbf{x}_1$.

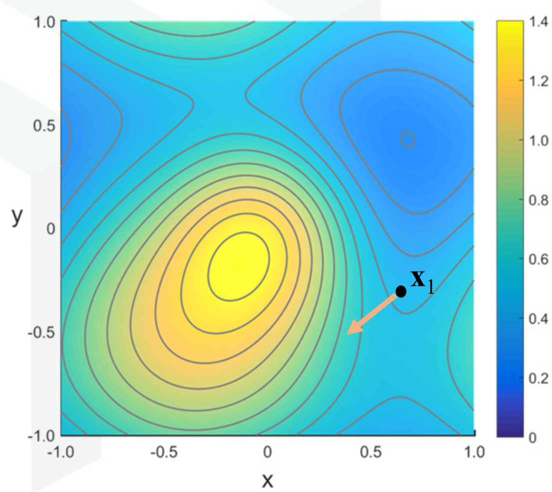


EMPossible

Concept of Steepest Ascent

Step 2 – Calculate the gradient because that points toward increasing values of $f(\mathbf{x}_1)$.

$$\nabla f\left(\mathbf{x}_1\right)$$

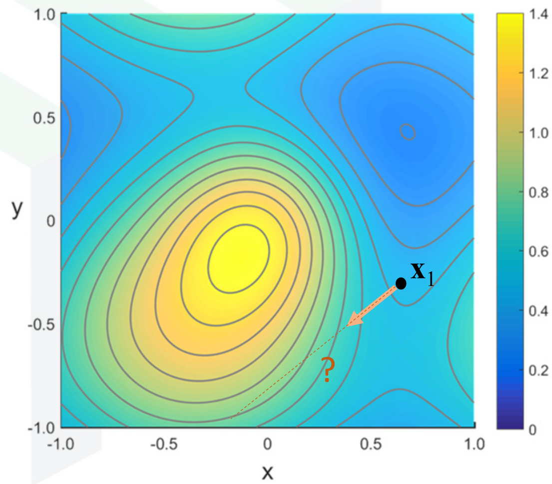


EMPossible

Concept of Steepest Ascent

Step 3 – Calculate next point $\mathbf{x}_2$ in direction of gradient. But how far?

$$\mathbf{x}_2 = \mathbf{x}_1 + ?$$



EMPossible

Concept of Steepest Ascent

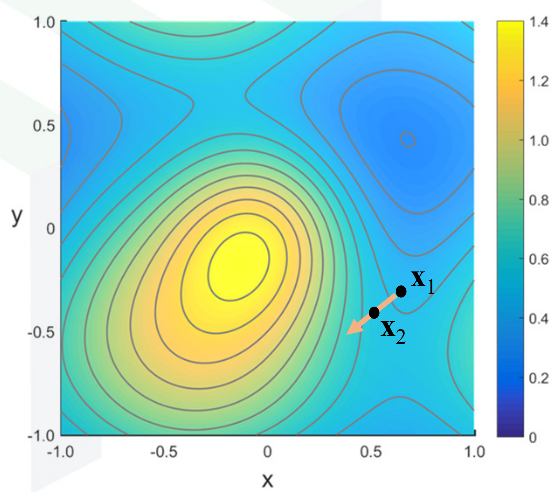
Step 3 – Calculate next point $\mathbf{x}_2$ in direction of gradient. But how far?

Moving too far along the gradient may cause the algorithm to go unstable and not find the maximum.

Moving not far enough will require many iterations to find the maximum.

For now, choose a constant $\alpha = 0.5$.

$$\mathbf{x}_2 = \mathbf{x}_1 + \alpha \nabla f\left(\mathbf{x}_1\right)$$

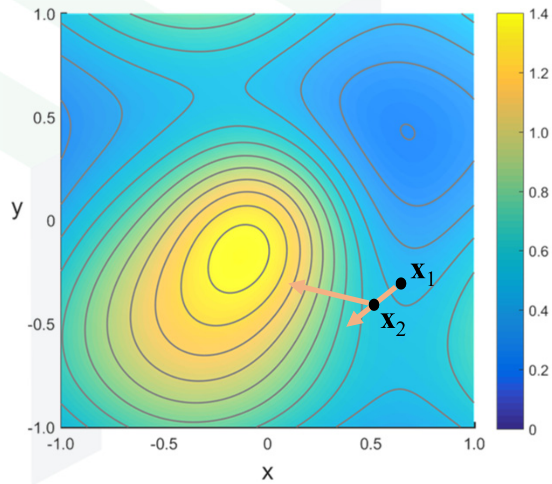


EMPossible

Concept of Steepest Ascent

Step 4 – Calculate the gradient at the second point $\mathbf{x}_2$.

$$\nabla f\left(\mathbf{x}_2\right)$$



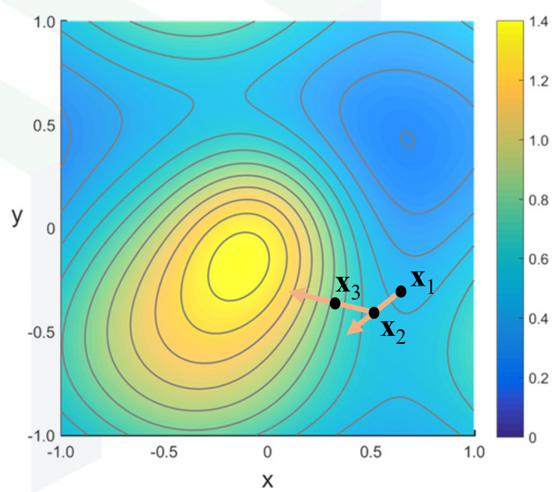
EMPossible

9

Concept of Steepest Ascent

Step 5 – Calculate the next point $\mathbf{x}_3$ along the gradient.

$$\mathbf{x}_3 = \mathbf{x}_2 + \alpha \nabla f\left(\mathbf{x}_2\right)$$



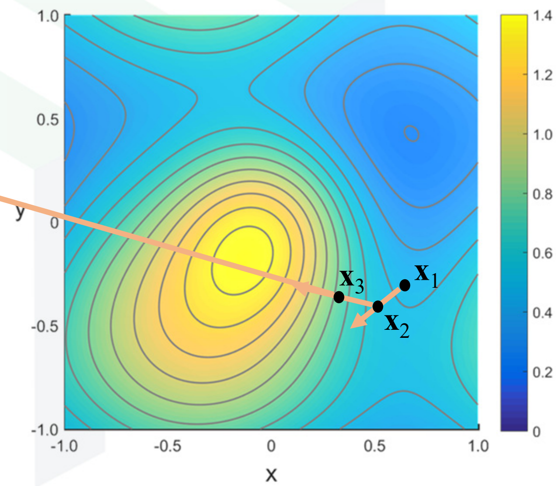
EMPossible

10

Concept of Steepest Ascent

Step 6 – Calculate the gradient at point $\mathbf{x}_3$.

$$\nabla f\left(\mathbf{x}_3\right)$$



The steep gradient caused the maximum to be overshot. The choice of α is important!

EMPossible

11

Summary of the Method

12

Steepest Ascent Method

It is desired to minimize the number of times the gradient is calculated. It is best to calculate the gradient once and then move in that direction until $f(x)$ stops increasing. At this point, the gradient is reevaluated and the process is repeated in the new direction.

Algorithm
1. Pick a starting point $\mathbf{x}$.
2. Calculate the gradient at this point: $\mathbf{g} = \nabla f(\mathbf{x})$
3. If the gradient is zero or less than some tolerance, we are done!
4. Otherwise, move in small increments in the direction of $\mathbf{g}$ until $f(\mathbf{x})$ stops increasing: $\mathbf{x} = \mathbf{x} + \alpha\mathbf{g}$
   Note: Think of this as searching along this direction like a 1D optimization and then efficiency can be improved greatly.
5. Go back to Step 2.

EMPossible

13

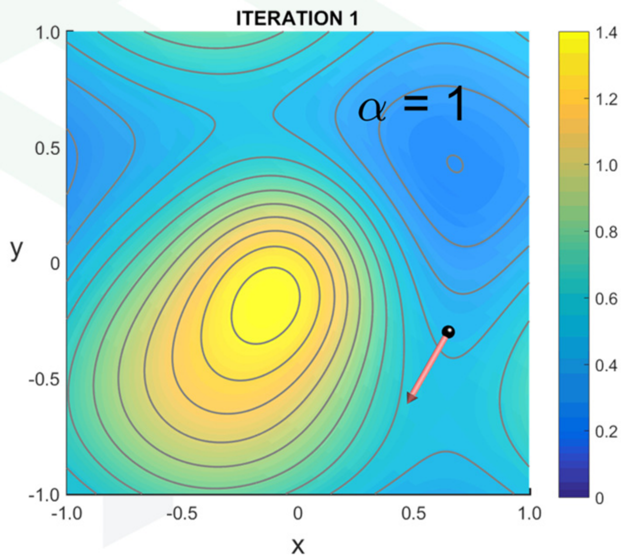
Choice of α

14

Choice of α (1 of 5)

Too large values of a can cause the algorithm to jump away from maximum.

At best, the algorithm converges on a different maximum.

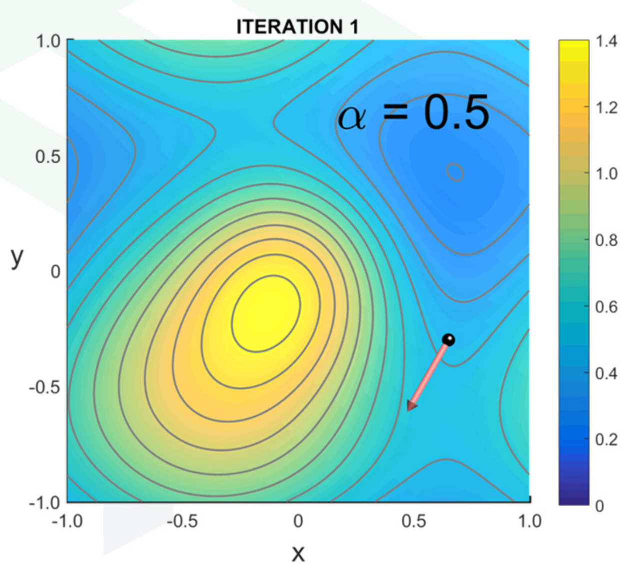


EMPossible

15

Choice of α (2 of 5)

Too large values of a can also cause the algorithm to oscillate about the maximum and never converge to it.

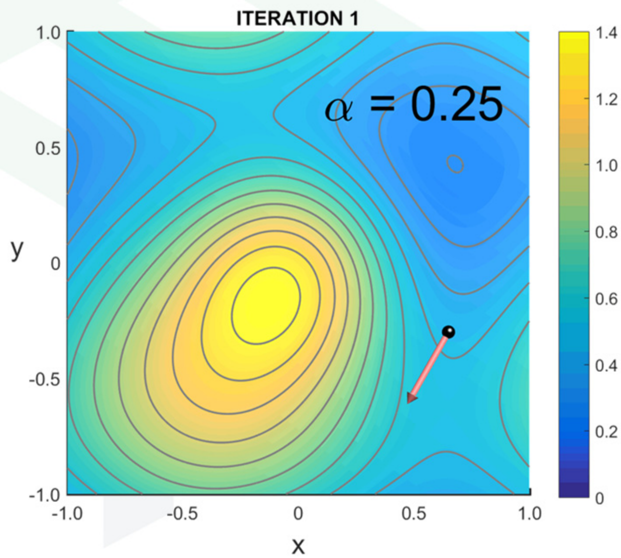


EMPossible

16

Choice of α (3 of 5)

For this example, $\alpha = 0.25$ seems like a very good choice. The best choice of α depends on the properties of the function. If the function varies wildly, choose small α. If the function is rather well behaved, larger values of α can converge faster.



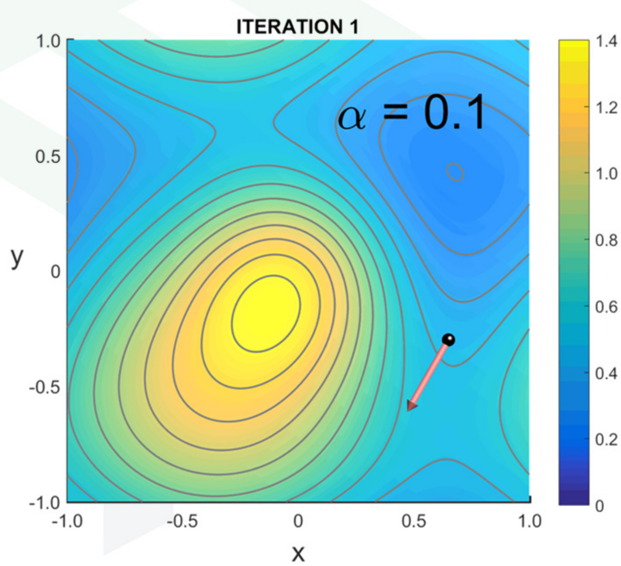
EMPossible

17

Choice of α (4 of 5)

$\alpha = 0.1$ seems like another good choice.

This is typically the value that I choose at first if nothing else is known.

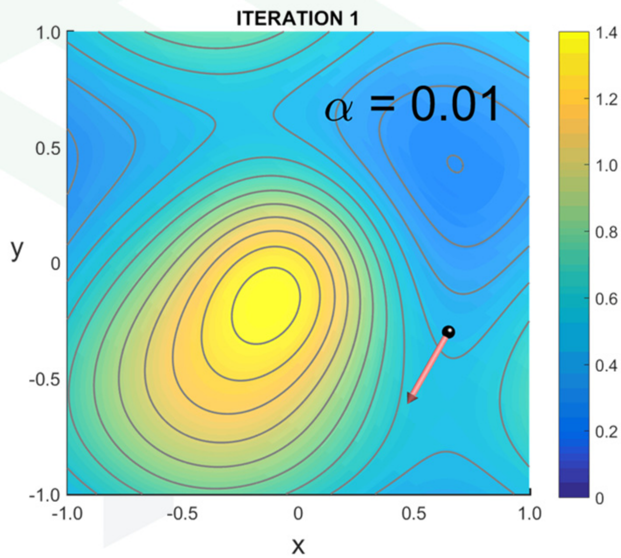


EMPossible

18

Choice of α (5 of 5)

Small values of α converge very slowly. This can be costly when evaluating the function is slow.



EMPossible

Example

Example (1 of 5)

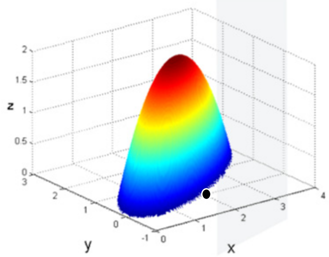
Problem

Find the maximum of the following function.

$$f(x,y) = 2xy + 2x - x^2 - 2y^2 \qquad \begin{array}{c} 0 < x < 4 \\ -0.5 < y < 2.5 \end{array}$$

Solution

Step 1 – Make an initial guess at the position.



$$\Rightarrow \quad \begin{array}{l} x_1 = 2 \\ y_1 = 0 \end{array}$$

EMPossible

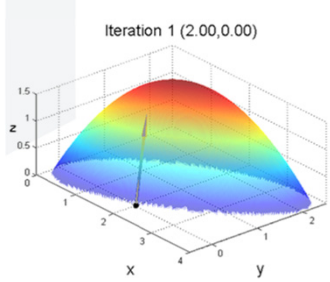
21

Example (2 of 5)

Step 2 – Calculate the gradient.

$$\begin{aligned} \nabla f(x,y) &= \frac{\partial f(x,y)}{\partial x}\hat{a}_x + \frac{\partial f(x,y)}{\partial y}\hat{a}_y \\ &= \frac{\partial}{\partial x}\left(2xy + 2x - x^2 - 2y^2\right)\hat{a}_x + \frac{\partial}{\partial y}\left(2xy + 2x - x^2 - 2y^2\right)\hat{a}_y \\ &= (2y + 2 - 2x)\hat{a}_x + (2x - 4y)\hat{a}_y \end{aligned}$$

$$\begin{aligned} \nabla f(2,0) &= [2\cdot 0 + 2 - 2\cdot 2]\hat{a}_x + [2\cdot 2 - 4\cdot 0]\hat{a}_y \\ &= -2\hat{a}_x + 4\hat{a}_y \end{aligned}$$



Step 3 – Gradient is not zero, so algorithm is not done.

EMPossible

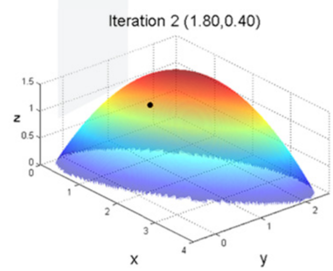
22

Example (3 of 5)

Step 4 – Move in direction of gradient.

$$\mathbf{x}_1 = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \qquad \mathbf{g}_1 = \begin{bmatrix} -2 \\ 4 \end{bmatrix} \qquad \text{Choose } \alpha = 0.1$$

$$\mathbf{x}_2 = \mathbf{x}_1 + \alpha \mathbf{g}_1 = \begin{bmatrix} 2 \\ 0 \end{bmatrix} + 0.1 \begin{bmatrix} -2 \\ 4 \end{bmatrix} = \begin{bmatrix} 1.8 \\ 0.4 \end{bmatrix}$$



EMPossible

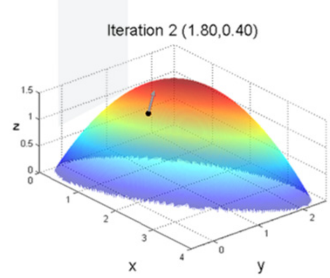
Example (4 of 5)

Step 5 – Go back to Step 2.

Step 2 – Calculate gradient at second point.

$$\mathbf{x}_2 = \begin{bmatrix} 1.8 \\ 0.4 \end{bmatrix} \qquad \mathbf{g}_2 = \begin{bmatrix} 2y + 2 - 2x \\ 2x - 4y \end{bmatrix} = \begin{bmatrix} 2 \cdot 0.4 + 2 - 2 \cdot 1.8 \\ 2 \cdot 1.8 - 4 \cdot 0.4 \end{bmatrix} = \begin{bmatrix} -0.8 \\ 2.0 \end{bmatrix}$$

Step 3 – Still not done!



EMPossible

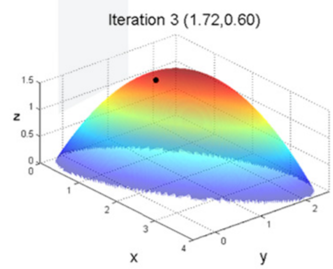
Example (5 of 5)

Step 5 – Go back to Step 2.

Step 4 – Calculate next point.

$$\mathbf{x}_2 = \begin{bmatrix} 1.8 \\ 0.4 \end{bmatrix} \qquad \mathbf{x}_3 = \mathbf{x}_2 + \alpha \mathbf{g}_2 = \begin{bmatrix} 1.8 \\ 0.4 \end{bmatrix} + 0.1 \begin{bmatrix} -0.8 \\ 2.0 \end{bmatrix} = \begin{bmatrix} 1.72 \\ 0.6 \end{bmatrix}$$

$$\mathbf{g}_2 = \begin{bmatrix} -0.8 \\ 2.0 \end{bmatrix}$$



EMPossible

25

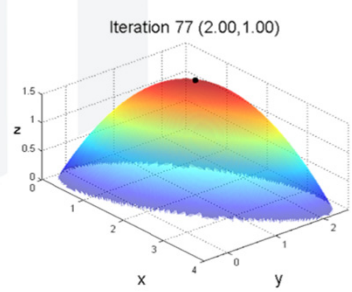
Example (6 of 6)

Step 5 – Go back to Step 2.

Step 5 – And so on…

After 77 iterations (tolerance 10^{-3}), the answer converges to

$$\mathbf{x}_e = \begin{bmatrix} 2.00 \\ 1.00 \end{bmatrix}$$



EMPossible

26