Open in Colab

- Download as webpdf:
  - jupyter nbconvert --to webpdf --allow-chromium-download chaps1_3.ipynb

Homework solutions

Quiz I

Suppose that $f''(c) = 0$, what are the sufficient conditions that c must furnish to be a relative minimum?

Answer: If $f''(c) = 0$, we must consider higher order terms in the expansion of equation (2.1.1). For example, we may expand up to third term and if $f'''(c) \neq 0$, its sign tells us about the behavior of the optimal point then. If the derivative is > 0, then $f(x)$ must have a relative maximum at c. If however, this higher-order derivative is < 0 $f(x)$ must have a relative minimum at c.

Specific to this question, $f^n(c) < 0$, where n is the order of the derivative.

Homework I

%%javascript MathJax.Hub.Config({ TeX: { equationNumbers: { autoNumber: "AMS" } } });

Prove that we have the *commutativity*, $\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{y} + \boldsymbol{x}$, and the associativity $\boldsymbol{x} + (\boldsymbol{y} + \boldsymbol{z}) = (\boldsymbol{x} + \boldsymbol{y}) + \boldsymbol{z}$

Solution:

Going by 3.1.3,

$$\boldsymbol{x}+\boldsymbol{y}=\begin{bmatrix} x_1+y_1 \\ x_2+y_2 \\ \vdots \\ x_N+y_N, \end{bmatrix} \tag{1}$$

It follows that,

$$\boldsymbol{y}+\boldsymbol{x}=\begin{bmatrix} y_1+x_1 \\ y_2+x_2 \\ \vdots \\ y_N+x_N, \end{bmatrix}=\begin{bmatrix} x_1+y_1 \\ x_2+y_2 \\ \vdots \\ x_N+y_N, \end{bmatrix} \tag{2}$$

We see that $\boldsymbol{x}+\boldsymbol{y}=\boldsymbol{y}+\boldsymbol{x}$, thus the addition of two vectors commute.

Associative proof follows the same logic as above. If you have problems showing that, please reach out to me.

Homework II

%%javascript MathJax.Hub.Config({ TeX: { equationNumbers: { autoNumber: "AMS" } } });

Prove that we have the *commutativity*, $\boldsymbol{x}-\boldsymbol{y}=\boldsymbol{y}-\boldsymbol{x}$, and the associativity $\boldsymbol{x}-(\boldsymbol{y}-\boldsymbol{z})=(\boldsymbol{x}-\boldsymbol{y})-\boldsymbol{z}$

Solution:

Again, we can extend (3.1.3) to subtraction as follows,

$$\boldsymbol{x}-\boldsymbol{y}=\begin{bmatrix} x_1-y_1 \\ x_2-y_2 \\ \vdots \\ x_N-y_N, \end{bmatrix} \tag{3}$$

It follows that,

$$\boldsymbol{y}-\boldsymbol{x}=\begin{bmatrix} y_1-x_1 \\ y_2-x_2 \\ \vdots \\ y_N-x_N, \end{bmatrix}=\begin{bmatrix} x_1-y_1 \\ x_2-y_2 \\ \vdots \\ x_N-y_N, \end{bmatrix} \tag{4}$$

We see that $\boldsymbol{x}-\boldsymbol{y}=\boldsymbol{y}-\boldsymbol{x}$, thus the addition of two vectors commute.

Associative proof follows the same logic as above. If you have problems showing that, please reach out to me.

Homework III

%%javascript MathJax.Hub.Config({ TeX: { equationNumbers: { autoNumber: "AMS" } } });

Prove that $\langle a\boldsymbol{x}+b\boldsymbol{y}, a\boldsymbol{x}+b\boldsymbol{y}\rangle = a^2\langle \boldsymbol{x},\boldsymbol{x}\rangle + 2ab\langle \boldsymbol{x},\boldsymbol{y}\rangle + b^2\langle \boldsymbol{y},\boldsymbol{y}\rangle$ is a non-negative quadratic form in the scalar variables a and b if $\boldsymbol{x}$ and $\boldsymbol{y}$ are real.

Solution:

We know from (3.1.6)[a-c] that

$$\langle \boldsymbol{x},\boldsymbol{y}\rangle = \langle \boldsymbol{y},\boldsymbol{x}\rangle \tag{5}$$
$$\langle \boldsymbol{x}+\boldsymbol{y},\boldsymbol{u}+\boldsymbol{v}\rangle = \langle \boldsymbol{x},\boldsymbol{u}\rangle + \langle \boldsymbol{x},\boldsymbol{v}\rangle + \langle \boldsymbol{y},\boldsymbol{u}\rangle + \langle \boldsymbol{y},\boldsymbol{v}\rangle \tag{6}$$
$$\langle c_1\boldsymbol{x},\boldsymbol{y}\rangle = c_1\langle \boldsymbol{x},\boldsymbol{y}\rangle \tag{7}$$

Therefore, it follows from (3.1.6b) that

$$\langle a\boldsymbol{x}+b\boldsymbol{y}, a\boldsymbol{x}+b\boldsymbol{y}\rangle = \langle a\boldsymbol{x},a\boldsymbol{x}\rangle + \langle a\boldsymbol{x},b\boldsymbol{y}\rangle + \langle b\boldsymbol{y},a\boldsymbol{x}\rangle + \langle b\boldsymbol{y},b\boldsymbol{y}\rangle \tag{8}$$

Furthermore, from (3.1.6c), we can further expand the above as

$$\begin{aligned}\langle a\boldsymbol{x}+b\boldsymbol{y}, a\boldsymbol{x}+b\boldsymbol{y}\rangle &= a\langle \boldsymbol{x},a\boldsymbol{x}\rangle + a\langle \boldsymbol{x},b\boldsymbol{y}\rangle + b\langle \boldsymbol{y},a\boldsymbol{x}\rangle + b\langle \boldsymbol{y},b\boldsymbol{y}\rangle && (9)\\ &= a\langle a\boldsymbol{x},\boldsymbol{x}\rangle + a\langle b\boldsymbol{y},\boldsymbol{x}\rangle + b\langle a\boldsymbol{x},\boldsymbol{y}\rangle + b\langle b\boldsymbol{y},\boldsymbol{y}\rangle && (10)\end{aligned}$$

where the last line follows from the commutativity property in (3.1.6a), so that we can write (following the logic in (3.1.6c)):

$$\begin{aligned}\langle a\boldsymbol{x}+b\boldsymbol{y}, a\boldsymbol{x}+b\boldsymbol{y}\rangle &= a^2\langle \boldsymbol{x},\boldsymbol{x}\rangle + \underbrace{ab\langle \boldsymbol{y},\boldsymbol{x}\rangle}_{\equiv ab\langle \boldsymbol{x},\boldsymbol{y}\rangle} + \underbrace{ba\langle \boldsymbol{x},\boldsymbol{y}\rangle}_{\equiv ab\langle \boldsymbol{x},\boldsymbol{y}\rangle} + b^2\langle \boldsymbol{y},\boldsymbol{y}\rangle && (11)\\ &= a^2\langle \boldsymbol{x},\boldsymbol{x}\rangle + ab\langle \boldsymbol{x},\boldsymbol{y}\rangle + ab\langle \boldsymbol{x},\boldsymbol{y}\rangle + b^2\langle \boldsymbol{y},\boldsymbol{y}\rangle && (12)\\ &= a^2\langle \boldsymbol{x},\boldsymbol{x}\rangle + 2ab\langle \boldsymbol{x},\boldsymbol{y}\rangle + b^2\langle \boldsymbol{y},\boldsymbol{y}\rangle && (13)\end{aligned}$$

Since a and b are scalar variables, and $\boldsymbol{x}$ and $\boldsymbol{y}$ are real, it follows that

$a^2\langle \boldsymbol{x},\boldsymbol{x}\rangle \geq 0 \quad \forall a \in \mathbb{R}$,

$b^2\langle \boldsymbol{y},\boldsymbol{y}\rangle \geq 0 \quad \forall b \in \mathbb{R}$, and

$2ab\langle \boldsymbol{x},\boldsymbol{y}\rangle \geq 0 \quad \forall (a,b) \in \mathbb{R}$

where $\mathbb{R}$ denotes the real line.

%%javascript MathJax.Hub.Config({ TeX: { equationNumbers: { autoNumber: "AMS" } } });

Homework IV

Hence, show that for real-valued vectors $\boldsymbol{x}$ and $\boldsymbol{y}$, that the Cauchy-Schwarz Inequality $\langle \boldsymbol{x},\boldsymbol{y}\rangle^2 \leq \langle \boldsymbol{x},\boldsymbol{x}\rangle\langle \boldsymbol{y},\boldsymbol{y}\rangle$ holds.

Solution There are many ways to prove this. I will give you one solution but if you are interested in other ways to prove this, please have a go at this document.

Expanding the left-hand-side of the above equation, we have

$\langle \boldsymbol{x},\boldsymbol{y}\rangle^2 = \langle \boldsymbol{x},\boldsymbol{y}\rangle \cdot \langle \boldsymbol{x},\boldsymbol{y}\rangle$

Suppose, we have $\boldsymbol{x} = [x_1, x_2, \cdots, x_n]^T$ and $\boldsymbol{y} = [y_1, y_2, \cdots, y_n]^T$, then we must have

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle^2 = \left(\sum_{i=1}^{n} x_i y_i \right)^2$$

Similarly, expanding the right-hand-side of the inequality, we have

$$\langle \boldsymbol{x}, \boldsymbol{x} \rangle \langle \boldsymbol{y}, \boldsymbol{y} \rangle = \sum_{i=1}^{n} x_i^2 \cdot \sum_{i=1}^{n} y_i^2$$

Subtracting, the problem translates to

or

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle^2 - \langle \boldsymbol{x}, \boldsymbol{x} \rangle \langle \boldsymbol{y}, \boldsymbol{y} \rangle \triangleq \left(\sum_{i=1}^{n} x_i y_i \right)^2 - \sum_{i=1}^{n} x_i^2 \cdot \sum_{i=1}^{n} y_i^2 \tag{14}$$

Seeing x and y are real, the first term i.e. $\left(\sum_{i=1}^{n} x_i y_i\right)^2$ is less than the second term i.e. $\sum_{i=1}^{n} x_i^2 \cdot \sum_{i=1}^{n} y_i^2$ so that together, they are nonpositive i.e.,

$$\left(\sum_{i=1}^{n} x_i y_i \right)^2 - \sum_{i=1}^{n} x_i^2 \cdot \sum_{i=1}^{n} y_i^2 \leq 0 \tag{15}$$

This is trivially shown as

$(x_1 y_1 + x_2 y_2 + \cdots + x_n y_n)^2 \leq (x_1^2 + x_2^2 + \cdots + x_n^2)(y_1^2 + y_2^2 + \cdots + y_n^2)$.

Therefore, we have

$$\left(\sum_{i=1}^{n} x_i y_i \right)^2 \leq \sum_{i=1}^{n} x_i^2 \cdot \sum_{i=1}^{n} y_i^2 \tag{16}$$

Hence, $\langle \boldsymbol{x}, \boldsymbol{y} \rangle^2 \leq \langle \boldsymbol{x}, \boldsymbol{x} \rangle \langle \boldsymbol{y}, \boldsymbol{y} \rangle$ holds.

Homework V

Using the result in Homework IV, show that for complex vectors x and y, $|\langle x, y \rangle|^2 \leq \langle x, \bar{x} \rangle \langle y, \bar{y} \rangle$

%%javascript MathJax.Hub.Config({ TeX: { equationNumbers: { autoNumber: "AMS" } } });

Solution V

Say $\bar{x} = a + ib$ and $\bar{y} = c + id$ where i is the imaginary number $\sqrt{-1}$. It follows that

$$\langle x, y \rangle = \langle a + ib, c + id \rangle \tag{17}$$

whose solution can be inferred from homework III as

$$\langle x, y \rangle = (ac - bd) + i(ad + bc). \tag{18}$$

Therefore,

$$|\langle x, y\rangle|^2 = (ac - bd)^2 - (ad + bc)^2 \tag{19}$$

Expanding the right hand side of the required (17) above, we have

$$\begin{aligned}\langle x, \bar{x}\rangle &= \langle a + ib, a - ib\rangle &(20)\\ &= \langle a, a\rangle - \langle a, ib\rangle + \langle ib, a\rangle - \langle ib, ib\rangle &(21)\\ &= |a|^2 - |b|^2; &(22)\end{aligned}$$

i.e. $\langle x, \bar{x}\rangle = |a|^2 - |b|^2$.

In a similar vein to (20), we can write,

$$\langle y, \bar{y}\rangle = |c|^2 - |d|^2. \tag{23}$$

Putting (19), (20), and (23) in the original equation (17), we find that

$$\begin{aligned}|\langle x, y\rangle|^2 &\triangleq (ac - bd)^2 - (ad + bc)^2 &(24)\\ &= \left[(ac)^2 + (bd)^2 - 2abcd\right] - \left[(ad)^2 + (bc)^2 + 2(abcd)\right] &(25)\\ &= (ac)^2 + (bd)^2 - 4abcd - (ad)^2 - (bc)^2. &(26)\end{aligned}$$

Now, following (20) and (23), we find that

$$\begin{aligned}\langle x, \bar{x}\rangle\langle y, \bar{y}\rangle &= (|a|^2 - |b|^2)(|c|^2 - |d|^2) &(27)\\ &= |a|^2|c|^2 - |a|^2|d|^2 - |b|^2|c|^2 + |b|^2|d|^2 &(28)\\ &= (ac)^2 + (bd)^2 + (ad)^2 + (bc)^2 &(29)\end{aligned}$$

Now, comparing (26) with (29), we find that

$$\begin{aligned}|\langle x, y\rangle|^2 &= \left[(ac)^2 + (bd)^2\right] - 4abcd - (ad)^2 - (bc)^2 &(30)\\ &= \left[\langle x, \bar{x}\rangle\langle y, \bar{y}\rangle - (ad)^2 - (bc)^2\right] - 4abcd - (ad)^2 - (bc)^2 &(31)\end{aligned}$$

It then becomes clear from above that

$$|\langle x, y\rangle|^2 \leq \langle x, \bar{x}\rangle\langle y, \bar{y}\rangle \tag{32}$$

holds, provided that $-(ad)^2 - (bc)^2 - 4abcd - (ad)^2 - (bc)^2 < 0$. Hence, proved.

Homework VI

Show that the *triangle inequality*

$$\langle \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{x} + \boldsymbol{y}\rangle^{\frac{1}{2}} \leq \langle \boldsymbol{x}, \boldsymbol{x}\rangle^{\frac{1}{2}} + \langle \boldsymbol{y}, \boldsymbol{y}\rangle^{\frac{1}{2}} \tag{33}$$

holds for any two real-valued variables.

Solution: The proof here is trivial. Simply replace $\boldsymbol{x}$ with a complex varible $\boldsymbol{x} = x_r + ix_g$ and follow the logic above, where the subscripts r and y respectively denote the real and imaginary parts.

The complex conjugate of $\boldsymbol{x}$ is defined as $\bar{\boldsymbol{x}} = x_r - ix_g$ <!-- By this logic, we find that

$$
\begin{aligned}
\langle \boldsymbol{x}+\boldsymbol{y}, \boldsymbol{x}+\boldsymbol{y} \rangle^{\frac{1}{2}} &= \langle x_r + ix_g + y_r + iy_g, x_r + ix_g + y_r + iy_g \rangle^{\frac{1}{2}} \\
&= \langle (x_r + y_r) + i\left(x_g + y_g\right), (x_r + y_r) + i\left(x_g + y_g\right) \rangle^{\frac{1}{2}} \\
&= \sqrt{\langle (x_r + y_r), (x_r + y_r) \rangle + \langle (x_r + y_r), i(x_g + y_g) \rangle + \langle i\left(x_g + y_g\right), (x_r + \\ + \langle i(x_g + y_g), i(x_g + y_g) \rangle} \\
&= \sqrt{\langle x_r + y_r \rangle^2 + \langle (x_r + y_r), i(x_g + y_g) \rangle + \langle i\left(x_g + y_g\right), (x_r + y_r) \rangle + \langle i(} \\
&= \sqrt{\langle x_r + y_r, x_r + y_r \rangle + 2\langle (x_r + y_r), i(x_g + y_g) \rangle - \langle (x_g + y_g), (x_g + y_g)}
\end{aligned}
$$

Going by the results of homework III, we find that the first term in the square root isis -->

Office Hours,

July 30, 2021

> Indented block

In [19]:
```python
import math
import numpy as np
```

In [20]:
```python
math.pi
```

Out[20]:
```
3.141592653589793
```

In [21]:
```python
np.pi
```

Out[21]:
```
3.141592653589793
```

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

In [22]:
```python
def quadratic(x, y):
    return x**2 + 2*x*y + y**2
```

In [23]:
```python
x = 25; y = 0.5
ans = quadratic(x, y)
print(ans)
```

```
650.25
```

Stationary Point

In [24]:
```python
def derivative(f, x):
    return 2*x + 2*y
```

In [25]:
```python
x = 4
y = 5
```

```
f = quadratic(x, y)
deri_x = derivative(f, 4)
print(deri_x)
```

18

In [26]:
```
def derivative_y(f, y):
    return 2*x + 2*y
```

Optimal Points

- Relative Maxima: Double Derivative of f with respect to the concerned variable (must be less than 0 i.e. < 0)
- Relative Miniima: Double Derivative of f with respect to the concerned variable + Relative Maxima: Double Derivative of f with respect to the concerned variable (must be less than 0 i.e. > 0)

Analytic Approach

minimize $Q(u,v) = au^2 + 2uvb + cv^2$

subject to

$u^2 + v^2 = 1$

minimize $Q(u,v) = au^2 + 2uvb + cv^2$

```
subject to
```

$u^2 + v^2 - 1 = 0$

Introduce the Lagrangian, λ, the optimization problem becomes

minimize $au^2 + 2uvb + cv^2 - \lambda(u^2 + v^2 - 1)$

Vectors and Matrices

$$Q(x_1, x_2, \ldots, x_N) = \sum_{i,j=1}^{N} a_{ij} x_i x_j \tag{39}$$

$$y_i = \sum_{j=1}^{N} a_{ij} x_j \qquad i = 1, 2, \ldots, N \tag{40}$$

$\boldsymbol{y} = A\boldsymbol{x}$

$$y_i = \sum_{j=1}^{N} a_{ij} x_j \qquad i = 1, 2, \ldots, N \tag{41}$$

$$y_1 = a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots \tag{42}$$
$$y_2 = a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots \tag{43}$$
$$y_3 = a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \cdots \tag{44}$$
$$\vdots = \dots \quad \dots \quad \dots \tag{45}$$
$$y_n = a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \cdots \tag{46}$$

Can be rewritten as

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13}x_3 & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23}x_3 & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33}x_3 & \cdots & a_{3n} \\ \vdots & \dots & \dots & \dots\vdots & \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} \tag{47}$$

Which can be easily written as

$$\underbrace{\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix}}_{\boldsymbol{y}} = \underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13}x_3 & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23}x_3 & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33}x_3 & \cdots & a_{3n} \\ \vdots & \dots & \dots & \dots\vdots & \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix}}_{\boldsymbol{A}} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}}_{\boldsymbol{x}} \tag{48}$$

or

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} \tag{49}$$

In [27]:

```
kortnie = [100, 200, 300, 400, 500]
ryan = [50, 60, 70, 80, 90]

def add(x, y):
    assert len(x)==len(y), 'two vectors must have equal dimensions before you can add.'

    z = []
    for i in range(len(x)):
        z_item = x[i]+y[i]
        z.append(z_item)

    return z

add(ryan, kortnie)
```

Out[27]:

```
[150, 260, 370, 480, 590]
```

In [28]:

```
import numpy as np

kortnie = np.array(([100, 200, 300, 4500, 500]))
ryan = np.array(([50, 60, 70, 80, 90]))

solution = ryan+kortnie

print(solution)
```

```
[ 150  260  370 4580  590]
```

Homework

Carry out an example of (3.1.4) in python and numpy using $c_1 = 0.5$ and $\boldsymbol{x}$ as ryan/kortnie vector above

Inner Products (or Dot Products)

In [29]:

```
x1 = np.random.randn(4, 1)
y1 = np.random.randn(4, 1)

print(x1)

print()

print(y1)
```

```
[[-0.33492539]
 [ 2.01204123]
 [ 0.59148775]
 [-0.90754677]]

[[0.91826913]
 [0.09763606]
 [0.94627913]
 [0.82928432]]
```

The dot product between x1 and y1

In [30]:

```
# for loop

z = 0

assert len(x1)==len(y1), 'x1 and x2 must be of the same size'

for i in range(len(x1)):
    z += (x1[i, 0]*y1[i, 0])
print(z)
```

```
-0.3040056770773464
```

In [31]:

```
x1.T.dot(y1)
```

Out[31]:

```
array([[-0.30400568]])
```

In [ ]:

```

```

In [ ]:

```

```