Maths for ML III

Nipun Batra

February 4, 2020

IIT Gandhinagar

Constrained Optimization

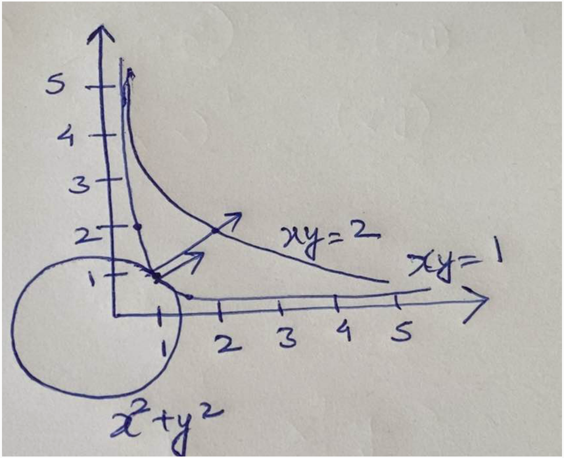
Extreme (max or min) $f(x, y) = x^2 + y^2$ s.t $xy = 1$

More generally Extrema f(x,...) s.t g(x,...) $= 0$

Constrained Optimization

Extreme (max or min) $f(x, y) = x^2 + y^2$ s.t $xy = 1$

More generally Extrema f(x,...) s.t g(x,...) $= 0$



Constrained Optimization

At extremum, (x^*, y^*), we get: $\nabla f(x^*, y^*) = \lambda \nabla g(x^*, y^*)$

Constrained Optimization

At extremum, (x^*, y^*), we get: $\nabla f(x^*, y^*) = \lambda \nabla g(x^*, y^*)$

$$\nabla f(x, y) = \begin{bmatrix} 2x \\ 2y \end{bmatrix} = \lambda \nabla g(x, y) = \lambda \begin{bmatrix} y \\ x \end{bmatrix}$$

Constrained Optimization

$$2x = \lambda y \tag{1}$$

$$2y = \lambda x \tag{2}$$

$$xy = 1 \tag{3}$$

Constrained optimization

We have three equations involving three variables. On solving the above equations, we get

$x = y = 1$

$\lambda = 2$

Constrained Optimization

Find extrema of $f(x, y) = x^2 + y^2$ s.t $x + y = 1$

Constrained Optimization

$$\nabla f(x, y) = \lambda \nabla g(x, y)$$

$$\nabla f(x, y) = \begin{bmatrix} 2x \\ 2y \end{bmatrix} \nabla g(x, y) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Constrained Optimization

$$2x = \lambda \tag{4}$$

$$2y = \lambda \tag{5}$$

$$x + y - 1 = 0 \tag{6}$$

On solving we get x = y = 0.5

Lagrangian Multiplier

For solving the form of equations: Extrema $f(.)$ s.t. $g(.) = 0$.

Lagrangian Multiplier

For solving the form of equations: Extrema $f(.)$ s.t. $g(.) = 0$.

For example, Maximize $f(x, y) = x^2 + y^2$ s.t. $x + y = 1$

Lagrangian Multiplier

For solving the form of equations: Extrema $f(.)$ s.t. $g(.) = 0$.

For example, Maximize $f(x, y) = x^2 + y^2$ s.t. $x + y = 1$

Construct a new function, Lagrangian
$L(x, y, \lambda) = f(x, y) + \lambda g(x, y)$ where λ is called the Lagrangian multiplier

Lagrangian Multiplier

For solving the form of equations: Extrema $f(.)$ s.t. $g(.) = 0$.

For example, Maximize $f(x, y) = x^2 + y^2$ s.t. $x + y = 1$

Construct a new function, Lagrangian
$L(x, y, \lambda) = f(x, y) + \lambda g(x, y)$ where λ is called the Lagrangian multiplier

- Set $\nabla L = 0$, i.e.

Lagrangian Multiplier

For solving the form of equations: Extrema $f(.)$ s.t. $g(.) = 0$.

For example, Maximize $f(x, y) = x^2 + y^2$ s.t. $x + y = 1$

Construct a new function, Lagrangian
$L(x, y, \lambda) = f(x, y) + \lambda g(x, y)$ where λ is called the Lagrangian multiplier

- Set $\nabla L = 0$, i.e.
- $\dfrac{\partial L}{\partial x} = 0$

Lagrangian Multiplier

For solving the form of equations: Extrema $f(.)$ s.t. $g(.) = 0$.

For example, Maximize $f(x, y) = x^2 + y^2$ s.t. $x + y = 1$

Construct a new function, Lagrangian
$L(x, y, \lambda) = f(x, y) + \lambda g(x, y)$ where λ is called the Lagrangian multiplier

- Set $\nabla L = 0$, i.e.
- $\dfrac{\partial L}{\partial x} = 0$
- $\dfrac{\partial L}{\partial y} = 0$

Lagrangian Multiplier

For solving the form of equations: Extrema $f(.)$ s.t. $g(.) = 0$.

For example, Maximize $f(x, y) = x^2 + y^2$ s.t. $x + y = 1$

Construct a new function, Lagrangian $L(x, y, \lambda) = f(x, y) + \lambda g(x, y)$ where λ is called the Lagrangian multiplier

- Set $\nabla L = 0$, i.e.
- $\dfrac{\partial L}{\partial x} = 0$
- $\dfrac{\partial L}{\partial y} = 0$
- $\dfrac{\partial L}{\partial \lambda} = 0$

Lagrangian Multiplier

Find the extrema of $f(x, y) = x^2 y$ s.t $g(x, y) = x^2 + y^2 = 1$

Lagrangian Multiplier

Find the extrema of $f(x, y) = x^2 y$ s.t $g(x, y) = x^2 + y^2 = 1$

$L(x, y, \lambda) = x^2 y + \lambda(x^2 + y^2 - 1)$

Lagrangian Multiplier

Find the extrema of $f(x,y) = x^2y$ s.t $g(x,y) = x^2 + y^2 = 1$

$L(x,y,\lambda) = x^2y + \lambda(x^2 + y^2 - 1)$

Compute the partial derivatives

Lagrangian Multiplier

$$\frac{\partial L}{\partial x} = 0 \implies 2xy + \lambda(2x) = 0 \tag{7}$$

$$\frac{\partial L}{\partial y} = 0 \implies x^2 + \lambda(2y) = 0 \tag{8}$$

$$\frac{\partial L}{\partial \lambda} = 0 \implies x^2 + y^2 - 1 = 0 \tag{9}$$

Case 1

x = 0

f(x,y) = 0

$y^2 = 1 \implies y = \pm 1$

$\lambda = 0$

Case 2

$x \neq 0 \implies y = -\lambda$

Case 2

$\mathrm{x} \neq 0 \implies y = -\lambda$

$x^2 = 2\lambda^2$

Case 2

$x \neq 0 \implies y = -\lambda$

$x^2 = 2\lambda^2$

Substitute the above values in Equation 9

Case 2

$x \neq 0 \implies y = -\lambda$

$x^2 = 2\lambda^2$

Substitute the above values in Equation 9

$$3\lambda^2 = 1 \implies \lambda = \pm\frac{1}{\sqrt{3}}$$

Case 2

$$\mathrm{x} \neq 0 \implies y = -\lambda$$

$$x^2 = 2\lambda^2$$

Substitute the above values in Equation 9

$$3\lambda^2 = 1 \implies \lambda = \pm\frac{1}{\sqrt{3}}$$

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

Max of $x^2y = \frac{2}{3}\sqrt{\frac{1}{3}}$

KKT Conditions

Used for constrained optimization of the form

Minimize $f(x)$, where $x \in \mathbb{R}^k$
such that

$$h_i(x) = 0, \ \forall i = 1, \ldots, m \text{ (m equalities)}$$
$$g_j(x) \leq 0, \ \forall j = 1, \ldots, n \text{ (n inequalities)}$$

Step 1

- Create a new function for minimization,

$$L(x, \lambda_1, \ldots, \lambda_m, \mu_1, \ldots, \mu_n) = f(x) + \sum_{i=1}^{m} \lambda_i h_i(x) + \sum_{j=1}^{n} \mu_j g_j(x)$$

where,

$\lambda_1 - \lambda_m$ are multipliers for the m equalities

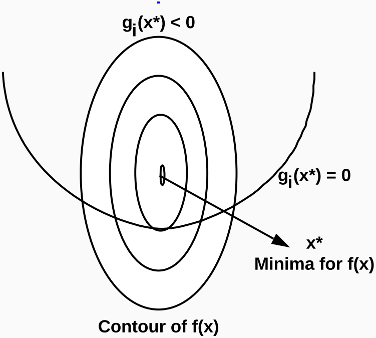
$\mu_1 - \mu_n$ are multipliers for the n inequalities

Step 2

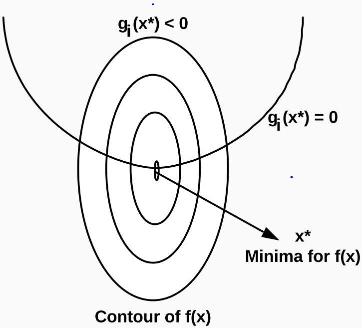
- Minimize $L(x, \lambda, \mu)$ w.rt. $x \implies \nabla_x L(x, \lambda, \mu) = 0$
  Gives k equations

Step 3

- Minimize $L(x, \lambda, \mu)$ w.rt. $\lambda \implies \nabla_\lambda L(x, \lambda, \mu) = 0$
  Gives m equations



$$g_i(x^*) \leq 0$$
$$\mu_i = 0$$



$$g_i(x^*) = 0$$

In both cases, $\mu_i g_i(x^*) = 0$

Minimize $f(x, y) = (x-4)^2 + (y-4)^2$

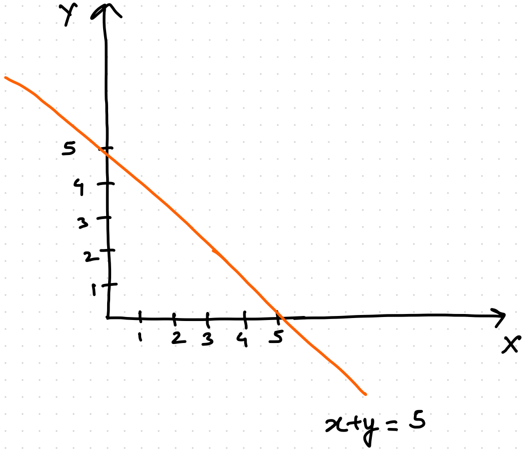
s.t. $x+y \geq 5$

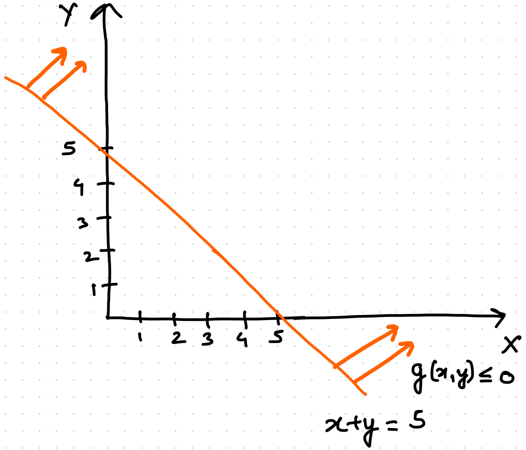
Minimize $f(x,y) = (x-4)^2 + (y-4)^2$

s.t. $x+y \geq 5$

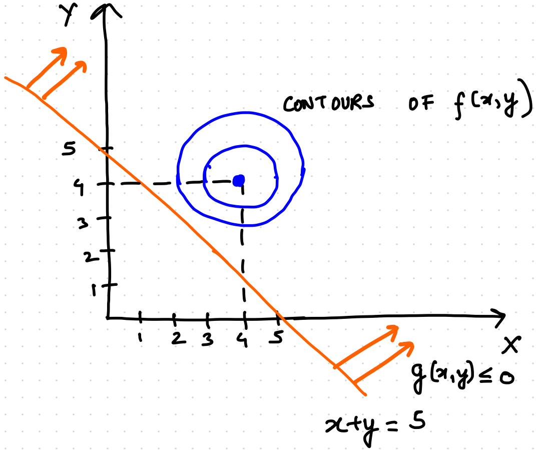
$f(x,y) = (x-4)^2 + (y-4)^2$

$g(x,y) = -x-y+5 \leq 0$

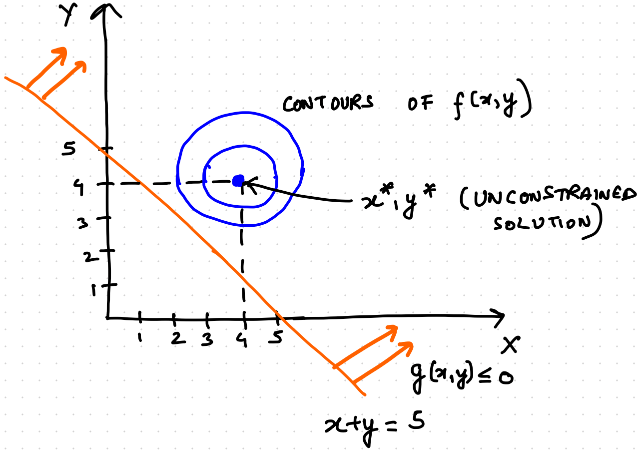

y
x
5
4
3
2
1
1
2
3
4
5
$x+y=5$



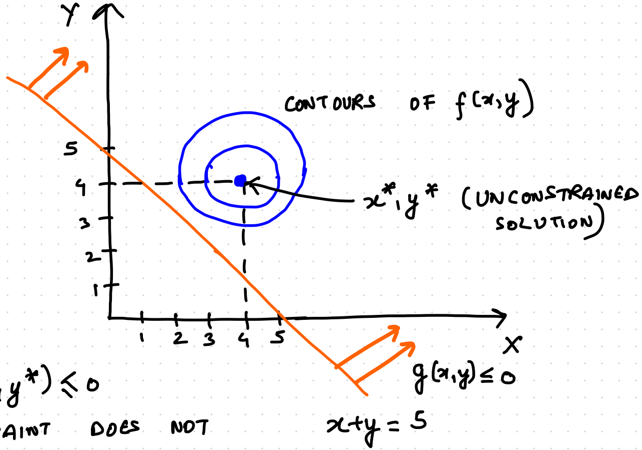
y
5
4
3
2
1
1
2
3
4
5
x
g(x,y) ≤ 0
x+y = 5



Y
X
5
4
3
2
1
1
2
3
4
5
CONTOURS OF f(x,y)
g(x,y) ≤ 0
x+y = 5



Y
X
5
4
3
2
1
1
2
3
4
5
CONTOURS OF f(x,y)
x*, y* (UNCONSTRAINED SOLUTION)
g(x,y) ≤ 0
x+y = 5



$g(x^*, y^*) \leq 0$

CONSTRAINT DOES NOT TAKE PART

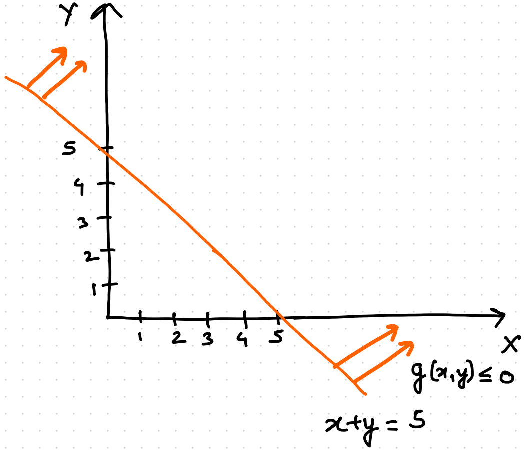
$\Rightarrow \mu = 0 \Rightarrow \boxed{\mu g(x,y) = 0}$

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

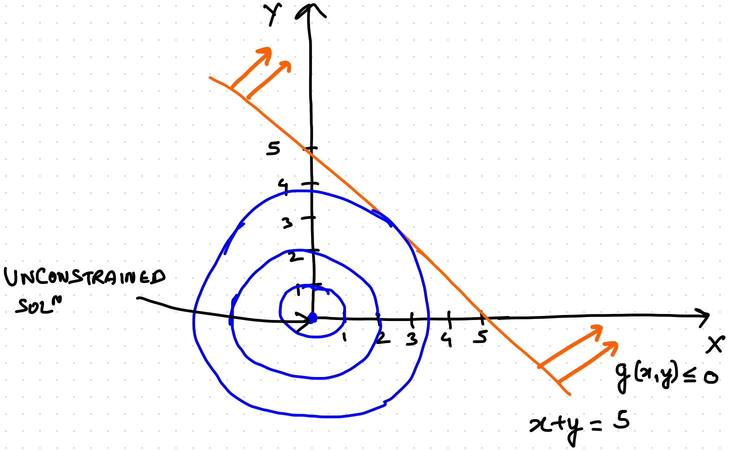
s.t. $x+y \geq 5$

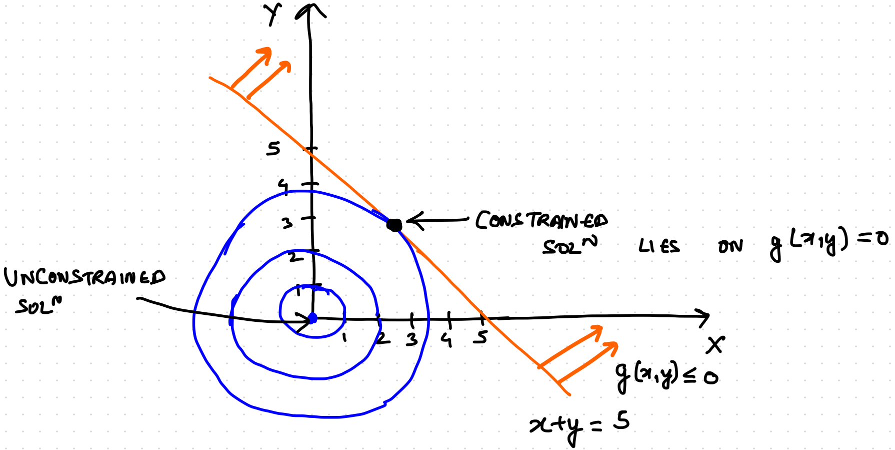
$f(x,y) = x^2 + y^2$

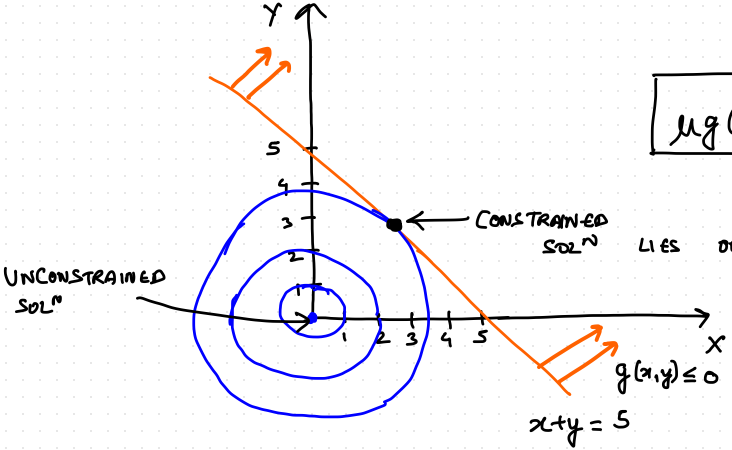
$g(x,y) = -x - y + 5 \leq 0$



y
5
4
3
2
1
1
2
3
4
5
x
g(x,y) ≤ 0
x+y = 5


Y
X
5
4
3
2
1
1
2
3
4
5
UNCONSTRAINED
SOLn
g(x,y) ≤ 0
x+y = 5


Y
X
5
4
3
2
1
1
2
3
4
5
CONSTRAINED
SOLN LIES ON g(x,y) = 0
UNCONSTRAINED
SOLN
g(x,y) ≤ 0
x+y = 5



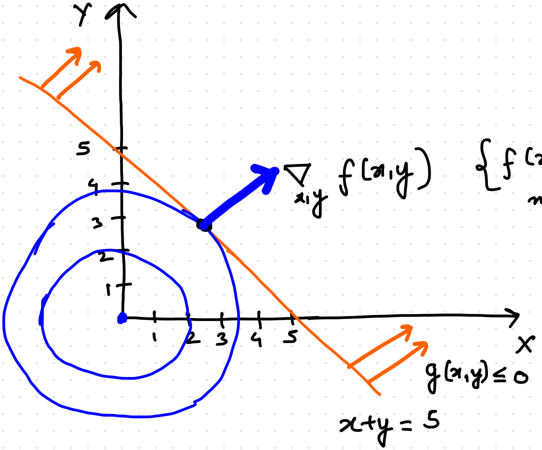
$$\boxed{\mu g(x,y)=0}$$

$\Uparrow$

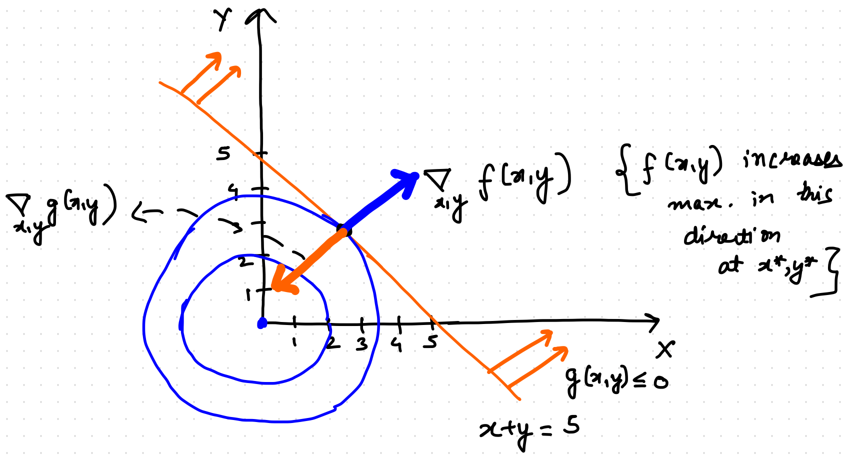
CONSTRAINED SOL^N LIES ON $g(x,y)=0$

WHY $\mu_i \geq 0 \quad \forall i$

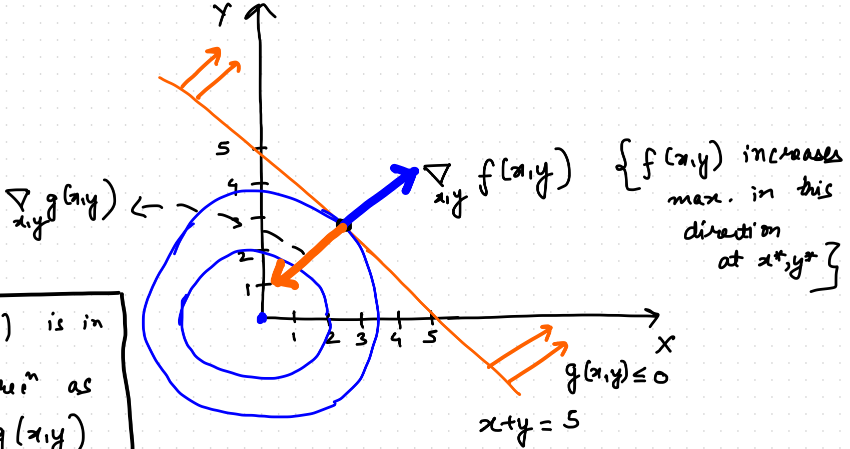
CONSIDER CASE WHEN $\mu \neq 0$



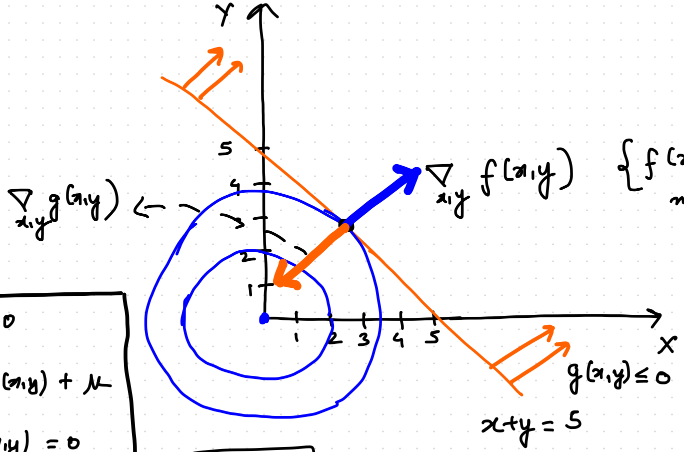
$\{f(x,y)$ increases max. in this direction at $x^*, y^*\}$



Y
X
5
4
3
2
1
1
2
3
4
5
$\nabla_{x,y}\, g(x,y)$
$\nabla_{x,y}\, f(x,y)$
{f(x,y) increases max. in this direction at x^*, y^*}
$g(x,y) \le 0$
$x+y=5$



$\nabla_{x,y} f(x,y)$ is in opp. direc^n as $\nabla_{x,y}\, g(x,y)$



{ $f(x,y)$ increases max. in this direction at x^*, y^* }

$\nabla_{x,y} L = 0$

$\Rightarrow \nabla_{x,y} f(x,y) + \mu \nabla_{x,y} g(x,y) = 0$

$\Rightarrow \mu = \dfrac{-\nabla f(x,y)}{\nabla g(x,y)} \Rightarrow \boxed{\mu > 0}$

KKT Conditions

Stationarity (For minimization)

$$\nabla_x f(x) + \sum_{i=1}^{m} \nabla_x \lambda_i h_i(x) + \sum_{i=1}^{n} \nabla_x \mu_i g_i(x) = 0$$

KKT Conditions

Stationarity (For minimization)

$$\nabla_x f(x) + \sum_{i=1}^{m} \nabla_x \lambda_i h_i(x) + \sum_{i=1}^{n} \nabla_x \mu_i g_i(x) = 0$$

Equality Constraints

$$\nabla_\lambda f(x) + \sum_{i=1}^{m} \nabla_\lambda \lambda_i h_i(x) + \sum_{i=1}^{n} \nabla_\lambda \mu_i g_i(x) = 0$$

$$\sum_{i=1}^{m} \nabla_\lambda \lambda_i h_i(x) = 0$$

KKT Conditions

Stationarity (For minimization)

$$\nabla_x f(x) + \sum_{i=1}^{m} \nabla_x \lambda_i h_i(x) + \sum_{i=1}^{n} \nabla_x \mu_i g_i(x) = 0$$

Equality Constraints

$$\nabla_\lambda f(x) + \sum_{i=1}^{m} \nabla_\lambda \lambda_i h_i(x) + \sum_{i=1}^{n} \nabla_\lambda \mu_i g_i(x) = 0$$

$$\sum_{i=1}^{m} \nabla_\lambda \lambda_i h_i(x) = 0$$

Inequality Constraints (Complementary Slackness)

$$\mu_i g_i(x) = 0 \forall i = 1, \dots, n$$

$$\mu_i \geq 0$$

Example

Minimize $x^2 + y^2$ such that,

$$x^2 + y^2 \leq 5$$
$$x + 2y = 4$$
$$x, y \geq 0$$

Example

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

Example

$$f(x, y) = x^2 + y^2$$
$$h(x, y) = x + 2y - 4$$

Example

$$f(x, y) = x^2 + y^2$$
$$h(x, y) = x + 2y - 4$$
$$g_1(x, y) = x^2 + y^2 - 5$$

Example

$$
\begin{aligned}
f(x, y) &= x^2 + y^2 \\
h(x, y) &= x + 2y - 4 \\
g_1(x, y) &= x^2 + y^2 - 5 \\
g_2(x, y) &= -x
\end{aligned}
$$

Example

$$
\begin{aligned}
f(x, y) &= x^2 + y^2 \\
h(x, y) &= x + 2y - 4 \\
g_1(x, y) &= x^2 + y^2 - 5 \\
g_2(x, y) &= -x \\
g_3(x, y) &= -y
\end{aligned}
$$

Example

$$f(x, y) = x^2 + y^2$$
$$h(x, y) = x + 2y - 4$$
$$g_1(x, y) = x^2 + y^2 - 5$$
$$g_2(x, y) = -x$$
$$g_3(x, y) = -y$$

$$L(x, y, \lambda, \mu_1, \mu_2, \mu_3) =$$
$$x^2 + y^2 + \lambda(x + 2y - 4) + \mu_1(x^2 + y^2 - 5) + \mu_2(-x) + \mu_3(-y)$$

Example

Stationarity

$\nabla_x L(x, y, \lambda, \mu_1, \mu_2, \mu_3) = 0$

$\implies 2x + \lambda + 2\mu_1 x - \mu_2 = 0 \ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$

$\nabla_y L(x, y, \lambda, \mu_1, \mu_2, \mu_3) = 0$

$\implies 2y + 2\lambda + 2\mu_1 y - \mu_3 = 0 \ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$

Example

Stationarity

$\nabla_x L(x, y, \lambda, \mu_1, \mu_2, \mu_3) = 0$

$\implies 2x + \lambda + 2\mu_1 x - \mu_2 = 0 \ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$

$\nabla_y L(x, y, \lambda, \mu_1, \mu_2, \mu_3) = 0$

$\implies 2y + 2\lambda + 2\mu_1 y - \mu_3 = 0 \ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$

Equality Constraint

$x + 2y = 4 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3)$

Example

Stationarity

$\nabla_x L(x, y, \lambda, \mu_1, \mu_2, \mu_3) = 0$

$\implies 2x + \lambda + 2\mu_1 x - \mu_2 = 0 \dotfill (1)$

$\nabla_y L(x, y, \lambda, \mu_1, \mu_2, \mu_3) = 0$

$\implies 2y + 2\lambda + 2\mu_1 y - \mu_3 = 0 \dotfill (2)$

Equality Constraint

$x + 2y = 4 \dotfill (3)$

Slackness

$\mu_1(x^2 + y^2 - 5) = 0 \dotfill (4)$

$\mu_2 x = 0 \dotfill (5)$

$\mu_3 y = 0 \dotfill (6)$

Example

From (6), $\mu_3 = 0$ or $y = 0$

But if, $y = 0$, then $x = 4$ according to (3) . This violates (1).

Hence, $y \neq 0$ and $\mu_3 = 0$

Example

From (6), $\mu_3 = 0$ or $y = 0$
But if, $y = 0$, then $x = 4$ according to (3) . This violates (1).
Hence, $y \neq 0$ and $\mu_3 = 0$

From (5), $\mu_1 = 0$ or $x = 0$
If $x = 0$, $y = 2$, which implies $x^2 + y^2 = 4 (\leq 5)$
Since (x,y) = (0,2) gives smaller $x^2 + y^2$ terms than 5,
Using (4), $\mu_1 = 0$

Example

From (6), $\mu_3 = 0$ or $y = 0$
But if, $y = 0$, then $x = 4$ according to (3) . This violates (1).
Hence, $y \neq 0$ and $\mu_3 = 0$

From (5), $\mu_1 = 0$ or $x = 0$
If $x = 0$, $y = 2$, which implies $x^2 + y^2 = 4(\leq 5)$
Since (x,y) = (0,2) gives smaller $x^2 + y^2$ terms than 5,
Using (4), $\mu_1 = 0$

On further solving we get,
$x = 0.8$
$y = 1.6$

MINIMIZE $x^2 + y^2$

s.t.

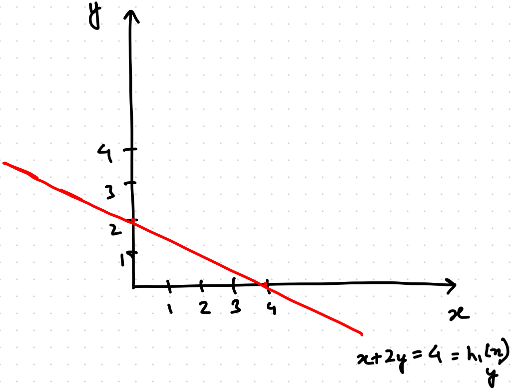
$x^2+y^2 \le 5$ or $g_1(x,y) = x^2+y^2-5 \le 0$ (μ_1)

$x+2y = 4$ or $h(x,y) = x+2y-4=0$

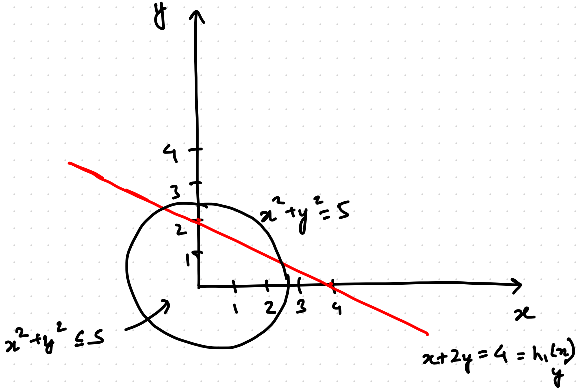
$x, y \geq 0$ or $g_2(x,y) = -x \le 0$ (μ_2)

$g_3(x,y) = -y \le 0$ (multiplier: μ_3)

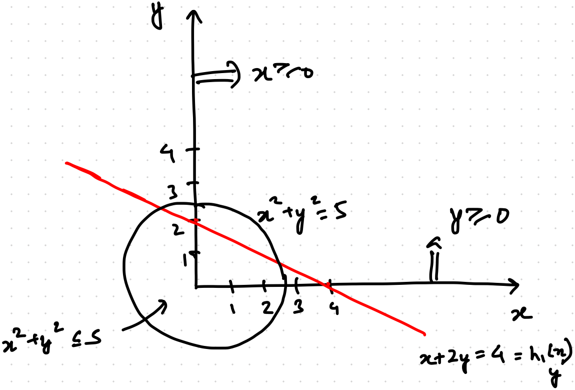
MINIMIZE $x^2 + y^2$



MINIMIZE $x^2 + y^2$

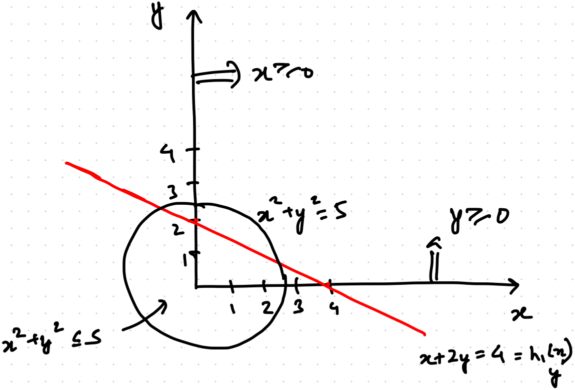


MINIMIZE x^2+y^2

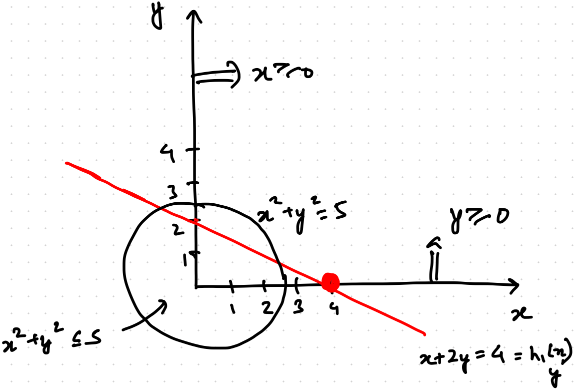


MINIMIZE $x^2 + y^2$

$\mu_3 = 0$ ¦ $\alpha > 0$



MINIMIZE $x^2 + y^2$



$\mu_3 = 0$ or > 0

$\mu_3 > 0$

$\Rightarrow y = 0$

$\Rightarrow x = 4$

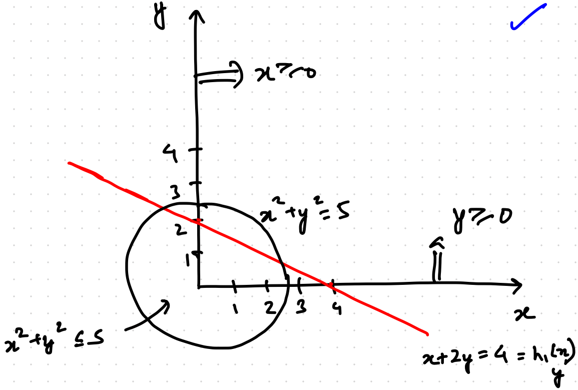
But doesn't satisfy

$x^2 + y^2 \leq 5$

$\therefore \mu_3 \not> 0$

MINIMIZE $x^2 + y^2$

$\mu_3 = 0$



$\mu_3 = 0$ ✓ or > 0

$\mu_3 > 0$

$\Rightarrow y = 0$

$\Rightarrow x = 4$

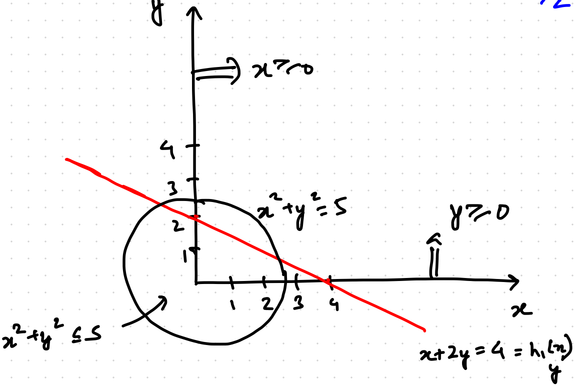
But doesn't satisfy $x^2+y^2 \leq 5$

$\therefore \mu_3 \not> 0$

MINIMIZE $x^2 + y^2$

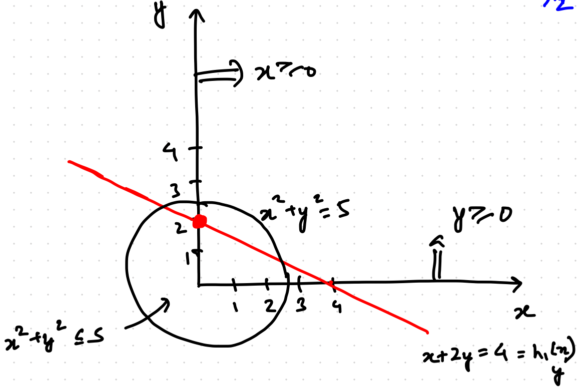
$\mu_3 = 0$

$\mu_2 = 0$ | $\mu_2 \geq 0$



MINIMIZE $x^2 + y^2$

$\mu_3 = 0$



$\mu_2 = 0$

$\mu_2 > 0$

$\Rightarrow x = 0$

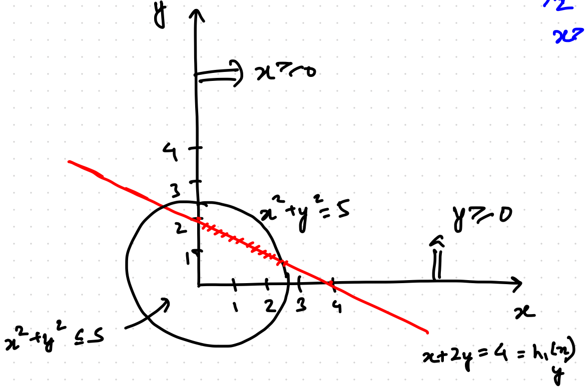
$\Rightarrow y = 2$

Feasible soln.

$x^2+y^2 = 4$

MINIMIZE $x^2 + y^2$

$\mu_3 = 0$



$\mu_2 = 0$

$x \geq 0$

$\mu_2 > 0$

$\Rightarrow x = 0$

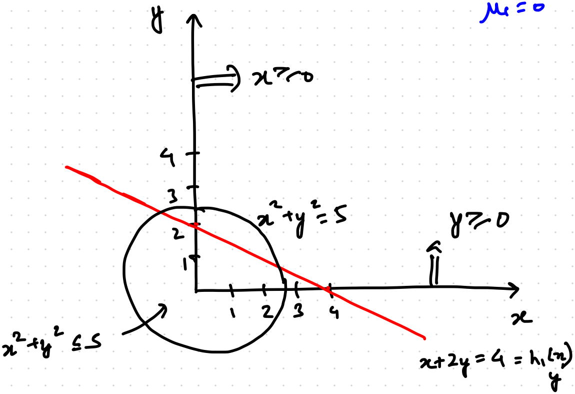
$\Rightarrow y = 2$

Feasible soln.

$x^2+y^2=4$

MINIMIZE $x^2 + y^2$

$\mu_3 = 0$



$\mu_1 = 0$

$\mu_1 > 0$

$\Rightarrow x^2+y^2 = 5$

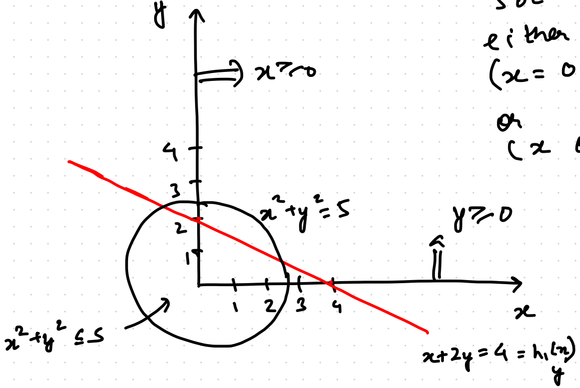
But we have seen better soln

$\therefore \mu_1 = 0$

MINIMIZE $x^2 + y^2$

$\mu_3 = 0$

$\mu_1 = 0$



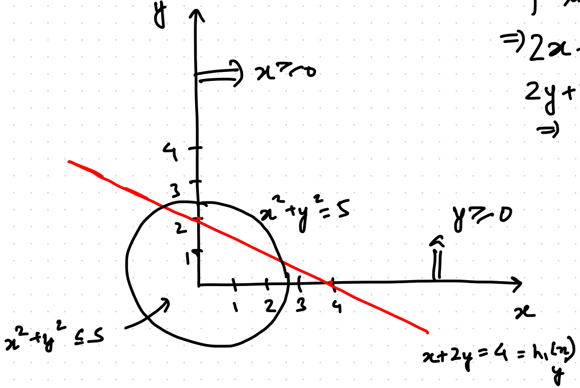
Sol^n is either $(x = 0, y = 2)$

or

(x on $x+2y=4$ line till it intersects circle)

MINIMIZE x^2+y^2

$\mu_3=0$
$\mu_1=0$
$\mu_2=0$



If $\mu_2=0$

$\Rightarrow 2x+\lambda=0$

$2y+2\lambda=0$

$\Rightarrow x=-\frac{\lambda}{2}$

$y=-\lambda$

$\therefore y=2x$

or

$x+2y=4$

$\Rightarrow 5x=4$

or $x=.8$

$y=1.6$