

Maths for ML III

Nipun Batra

December 26, 2023

IIT Gandhinagar

Constrained Optimization

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

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

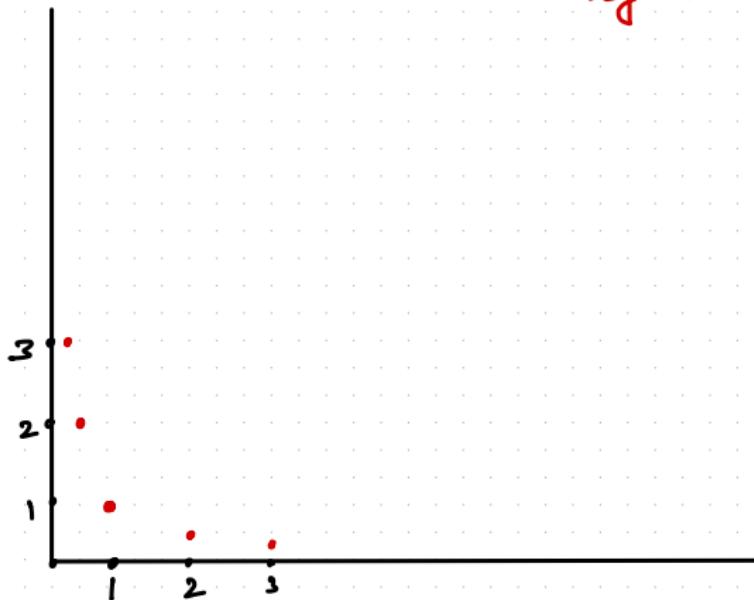
Constrained Optimization

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

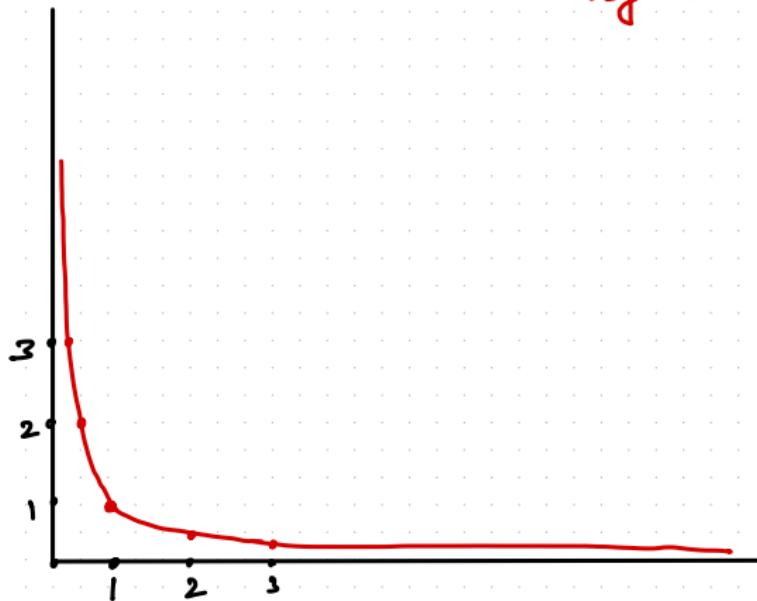
More generally Extrema $f(x, \dots)$ s.t $g(x, \dots) = 0$

Acknowledgments: Inspired by Khan Academy videos

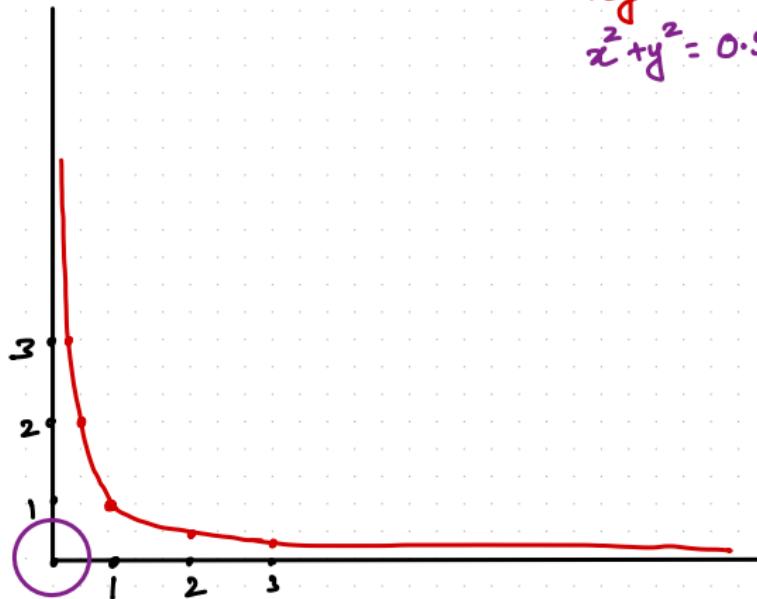
(DRAWN ONLY IN
 $xy = 1$ FIRST QUADRANT)



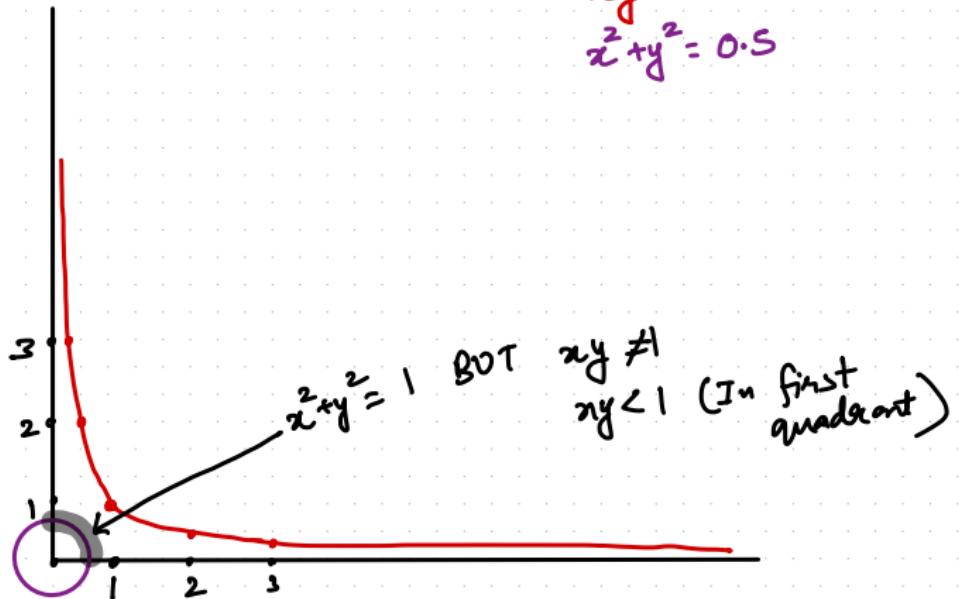
$$xy = 1$$



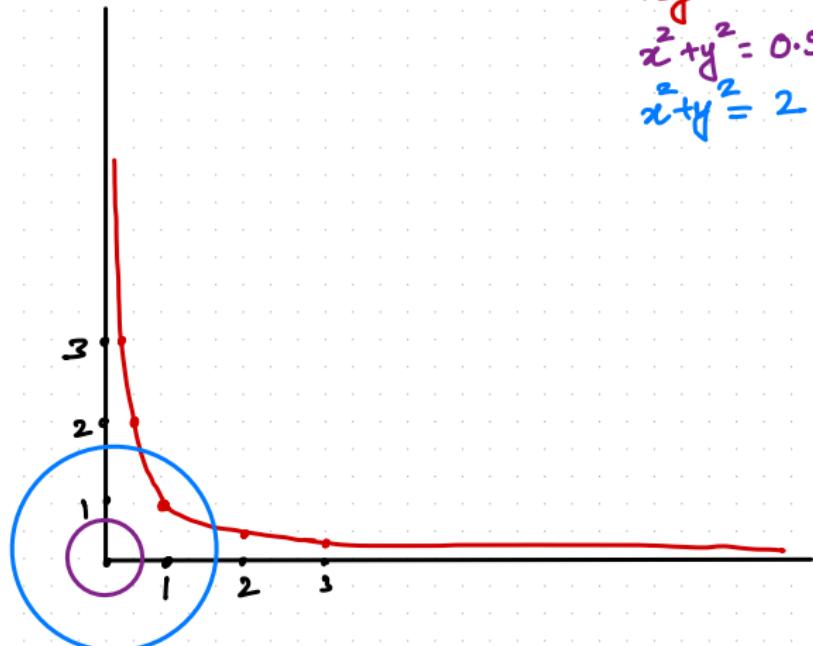
$$xy = 1$$
$$x^2 + y^2 = 0.5$$



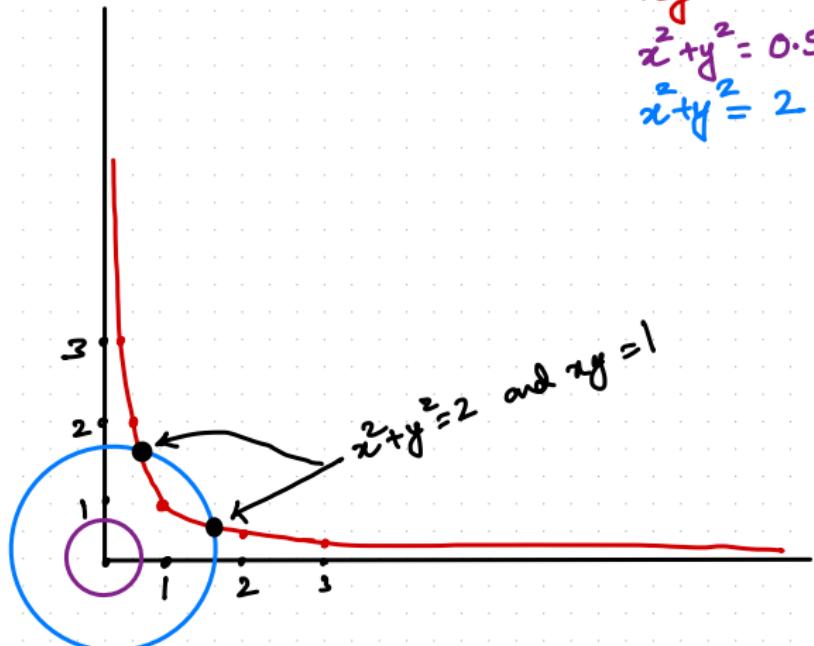
$$xy = 1$$
$$x^2 + y^2 = 0.5$$



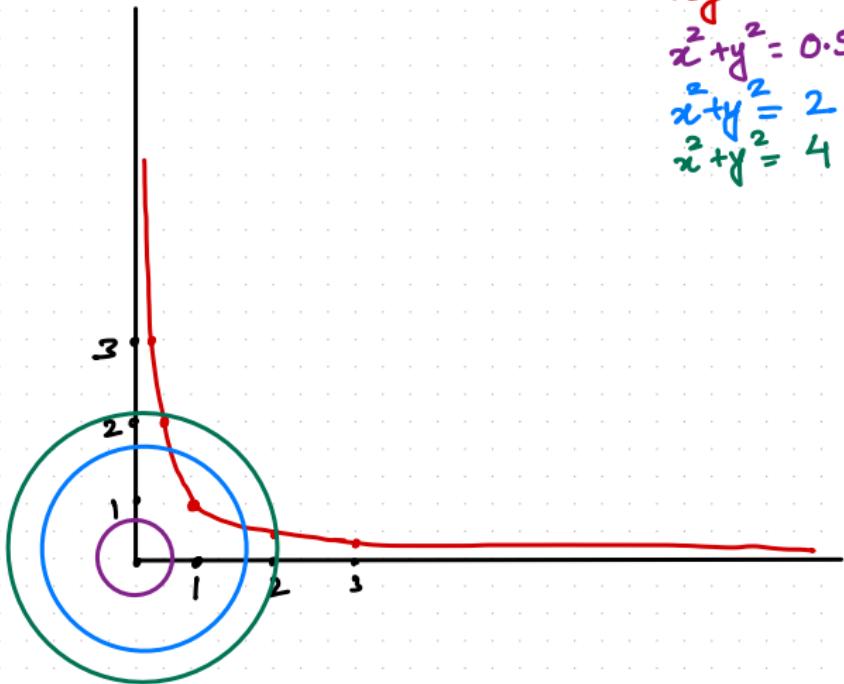
$$\begin{aligned}xy &= 1 \\x^2 + y^2 &= 0.5 \\x^2 + y^2 &= 2\end{aligned}$$



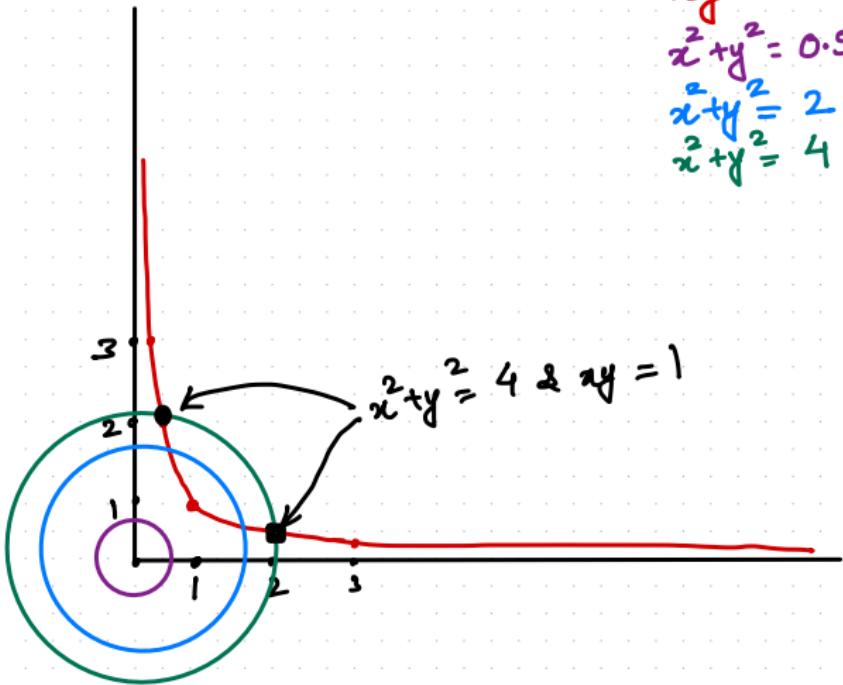
$$xy = 1$$
$$x^2 + y^2 = 0.5$$
$$x^2 + y^2 = 2$$



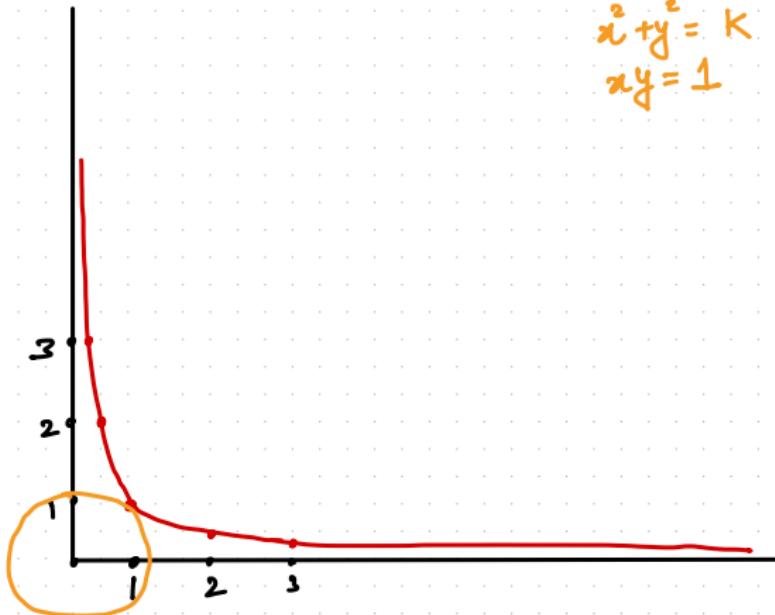
$$\begin{aligned}xy &= 1 \\x^2 + y^2 &= 0.5 \\x^2 + y^2 &= 2 \\x^2 + y^2 &= 4\end{aligned}$$



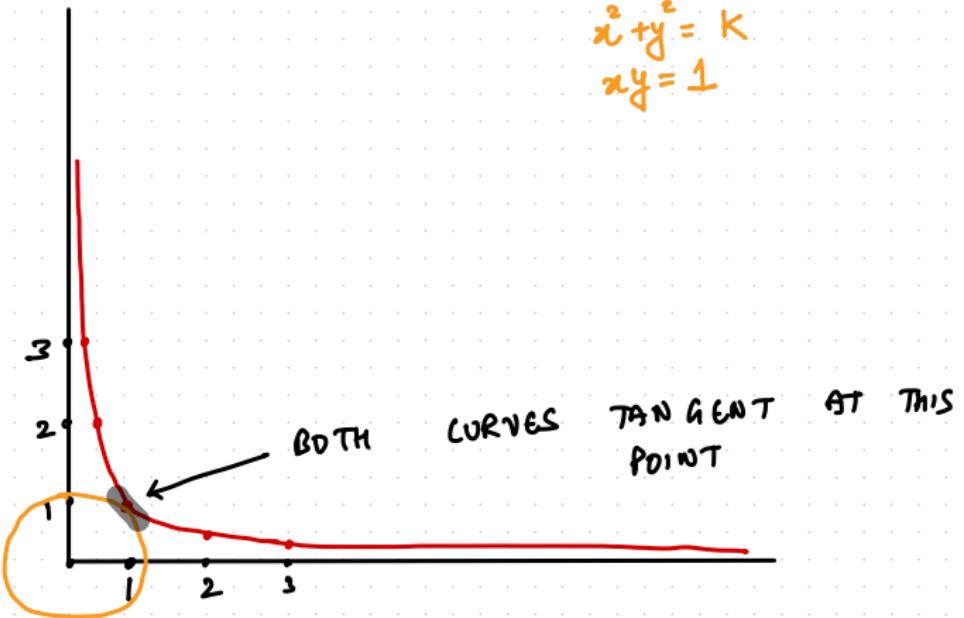
$$\begin{aligned}xy &= 1 \\x^2 + y^2 &= 0.5 \\x^2 + y^2 &= 2 \\x^2 + y^2 &= 4\end{aligned}$$



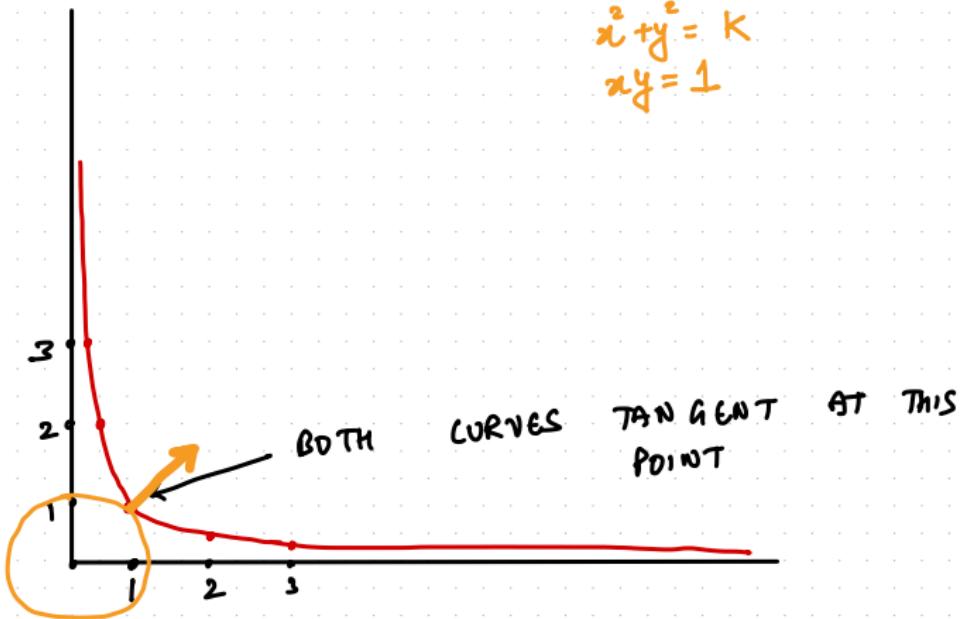
$$x^2 + y^2 = K$$
$$xy = 1$$



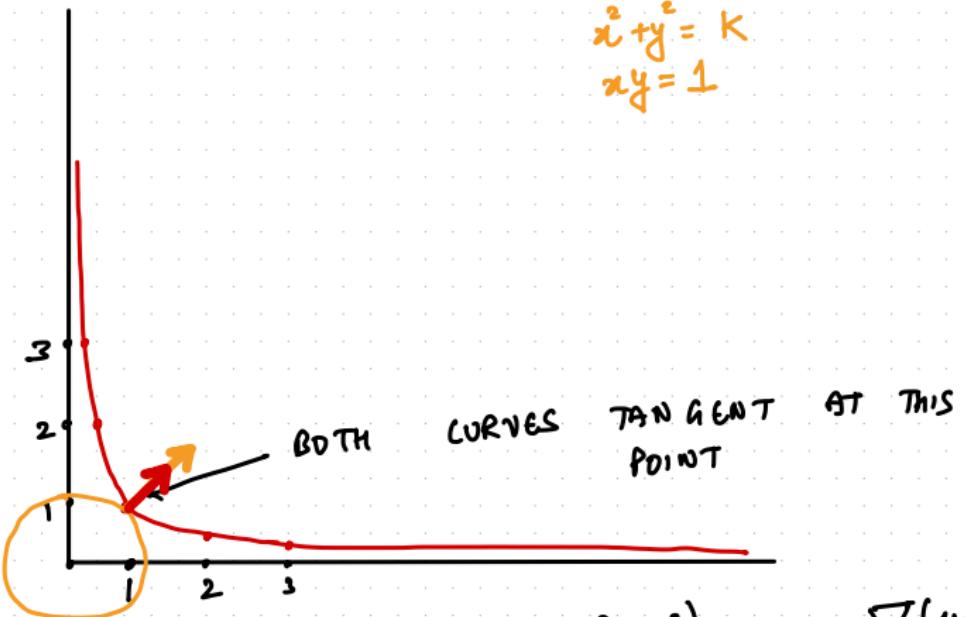
$$x^2 + y^2 = K$$
$$xy = 1$$



$$x^2 + y^2 = K$$
$$xy = 1$$



$$x^2 + y^2 = K$$
$$xy = 1$$



$$\nabla(x^2 + y^2)_{x^2, y^2} = \lambda \nabla(xy)_{x^*, y^*}$$

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

$$2x = \lambda y \quad (1)$$

$$2y = \lambda x \quad (2)$$

$$xy = 1 \quad (3)$$

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

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

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

Constrained Optimization

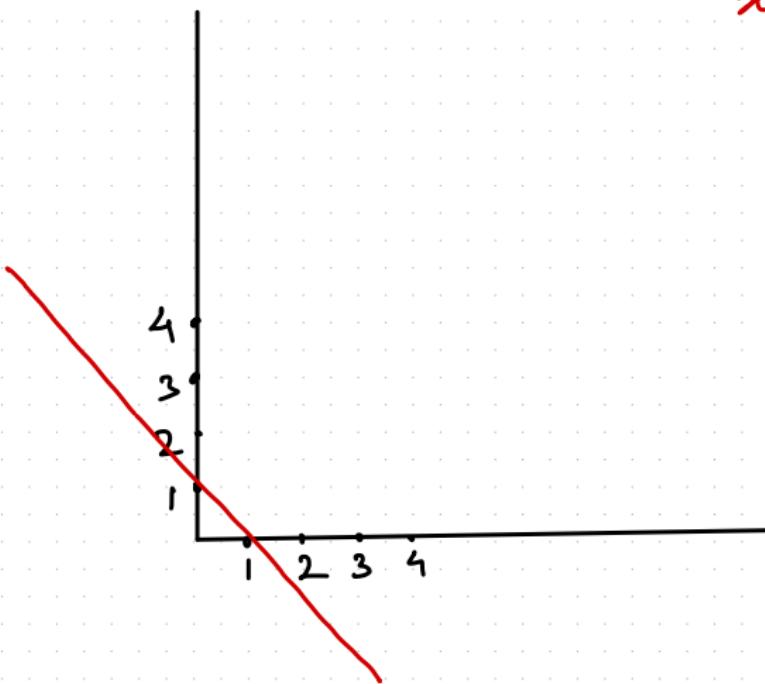
$$2x = \lambda \quad (4)$$

$$2y = \lambda \quad (5)$$

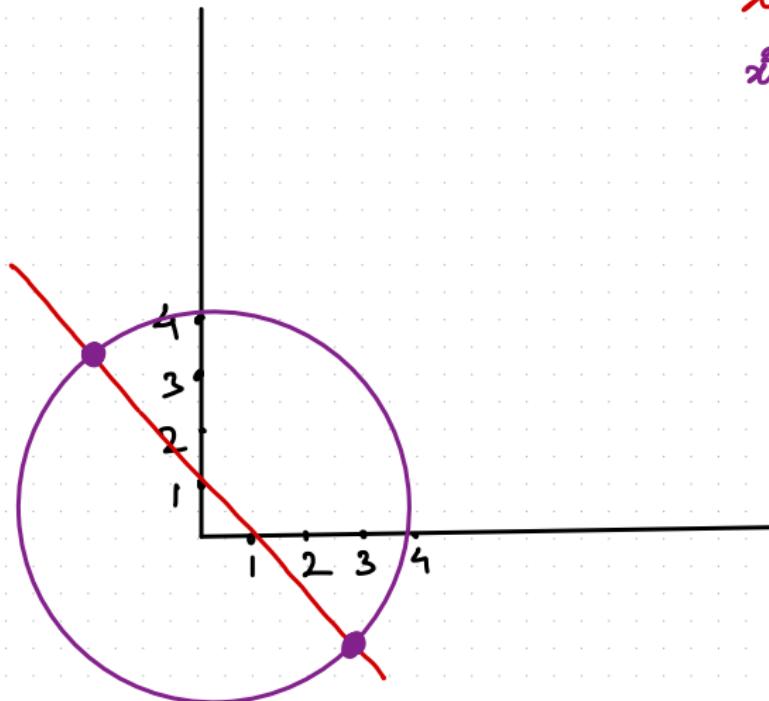
$$x + y - 1 = 0 \quad (6)$$

On solving we get $x = y = 0.5$

$$x+ty=1$$



$$x+y=1$$
$$x^2+y^2=4$$

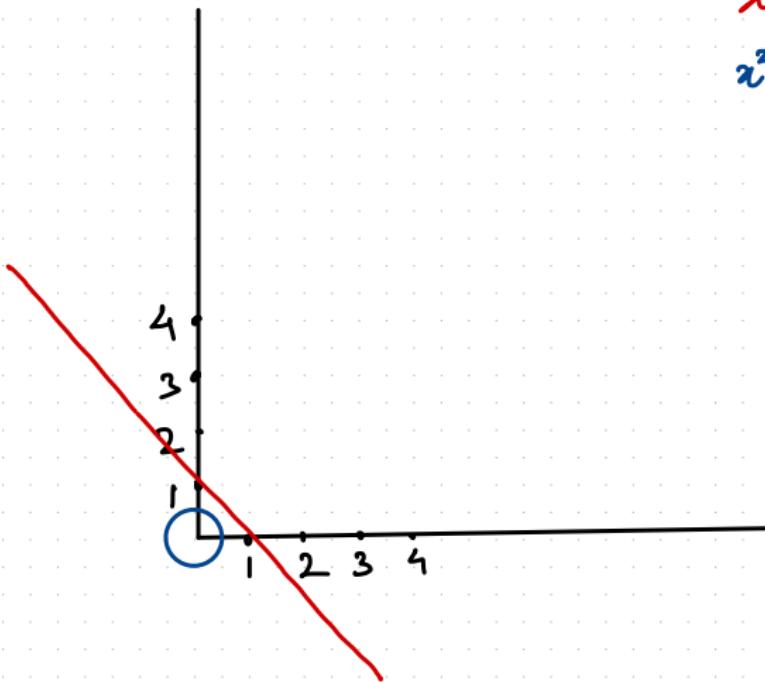


$$x+y=1$$

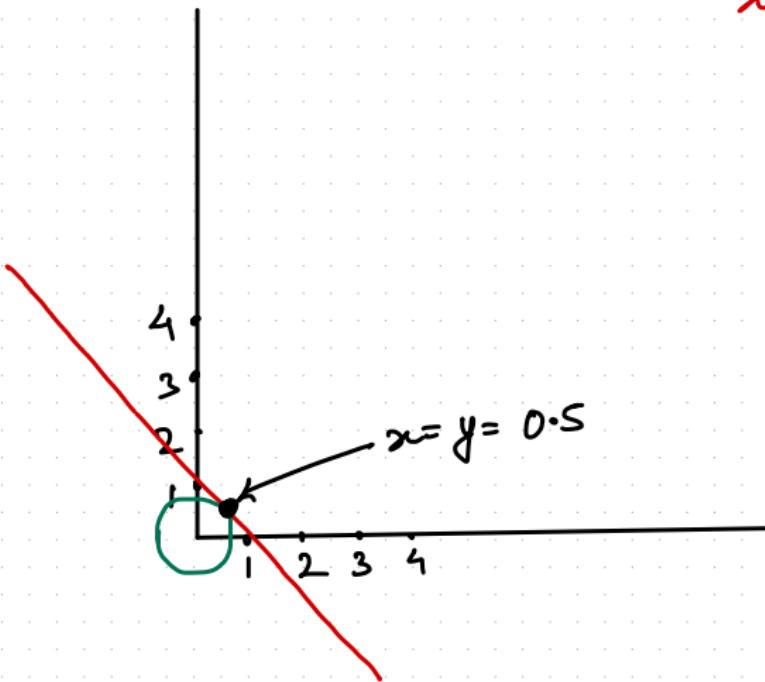
$$x^2+y^2 = K$$

But

$$x+y \leq 1$$



$$x+y=1$$



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 $f(\cdot)$ s.t. $g(\cdot) = 0$.

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 $f(\cdot)$ s.t. $g(\cdot) = 0$.

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.
- $\frac{\partial L}{\partial x} = 0$

Lagrangian Multiplier

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

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.

- $\frac{\partial L}{\partial x} = 0$

- $\frac{\partial L}{\partial y} = 0$

Lagrangian Multiplier

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

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.

- $\frac{\partial L}{\partial x} = 0$

- $\frac{\partial L}{\partial y} = 0$

- $\frac{\partial L}{\partial \lambda} = 0$

Lagrangian Multiplier

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

Lagrangian Multiplier

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

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

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

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

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

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

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

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

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

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

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

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

Substitute the above values in Equation 9

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

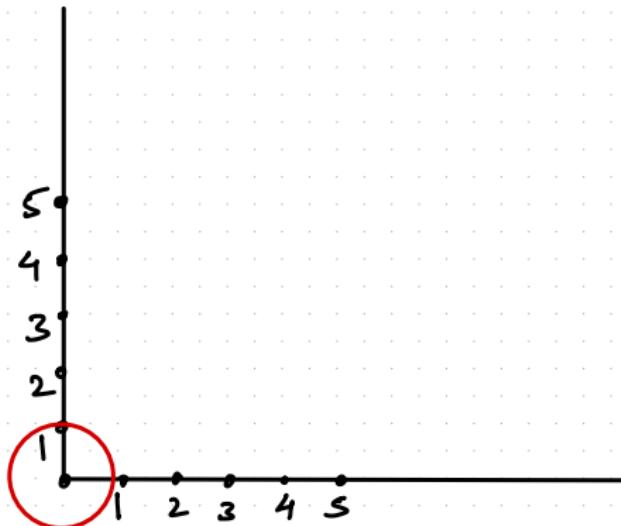
$$y = \pm \frac{1}{\sqrt{3}}$$

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

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

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

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



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

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

$$x^2y = 1$$

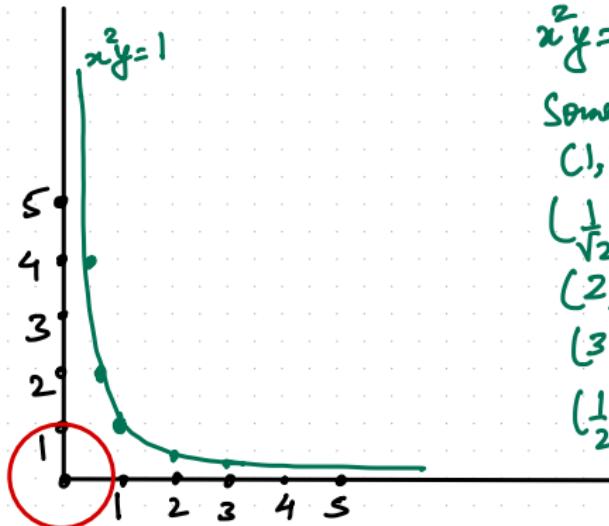
Some points
 $(1, 1)$

$$\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$$

$$\left(2, \frac{1}{4}\right)$$

$$\left(3, \frac{1}{9}\right)$$

$$\left(\frac{1}{2}, 4\right)$$



KKT Conditions

Used for constrained optimization of the form

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

$$h_i(x) = 0, \forall i = 1, \dots, m \text{ (m equalities)}$$

$$g_j(x) \leq 0, \forall j = 1, \dots, n \text{ (n inequalities)}$$

Step 1

- Create a new function for minimization,

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

where,

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

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

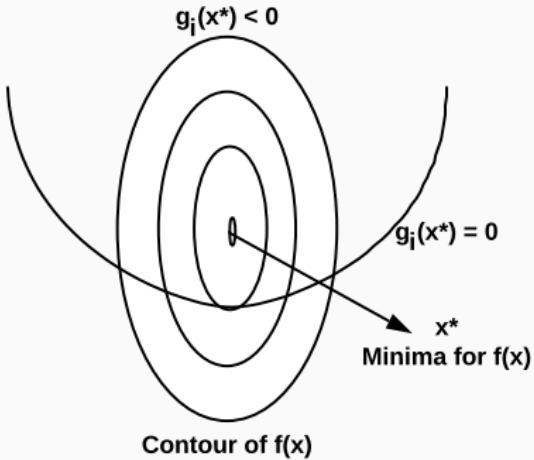
Step 2

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

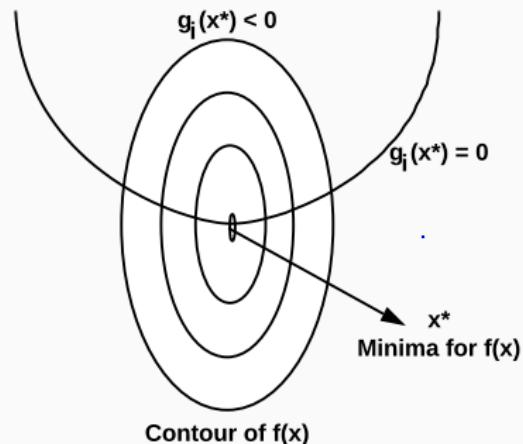
Step 3

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

Step 4



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



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

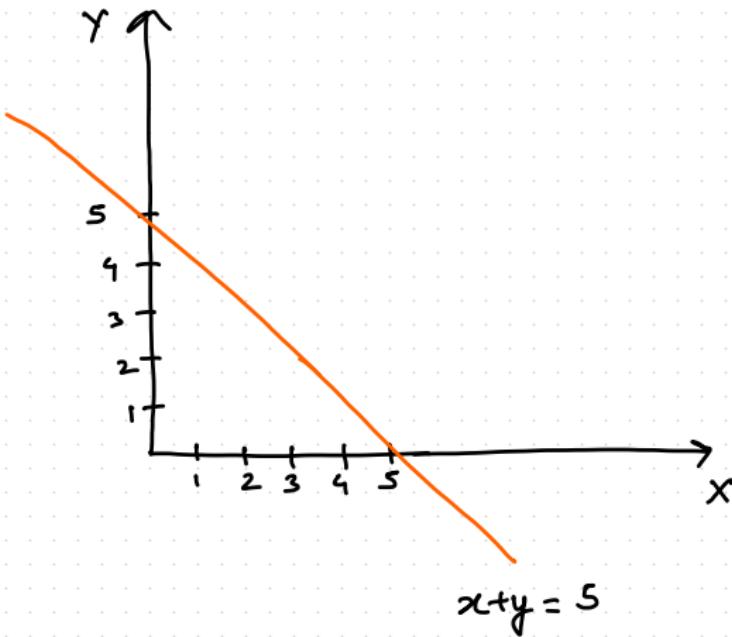
In both cases, $\mu_i g_i(x^*) = 0$

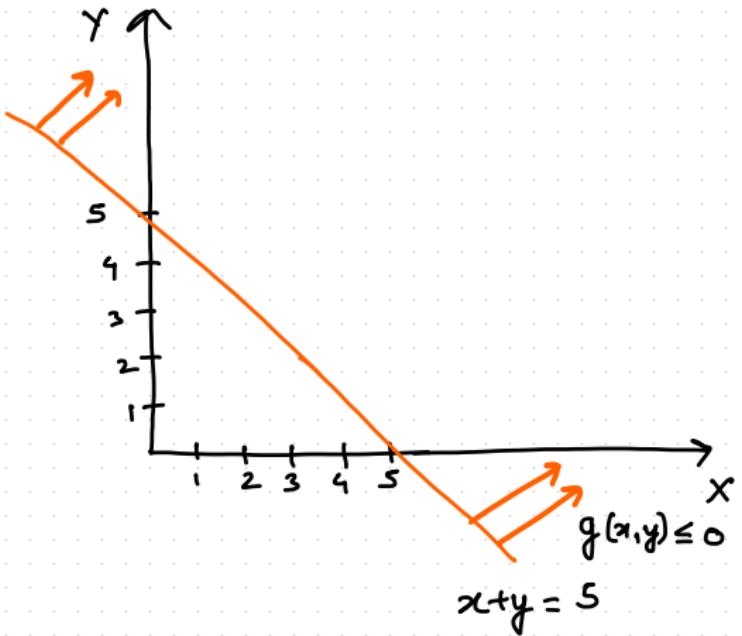
$$\text{Minimize } f(x, y) = (x-4)^2 + (y-4)^2$$

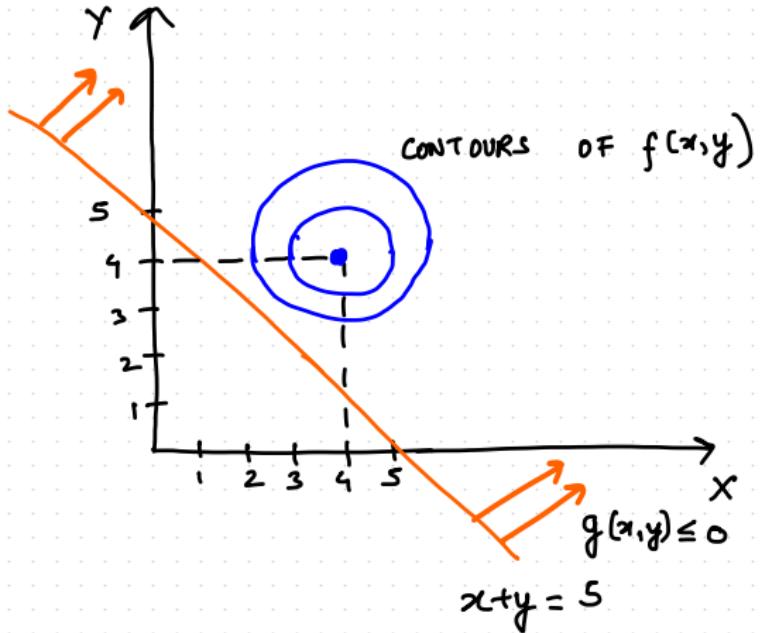
$$\text{s.t. } x+y \geq 5$$

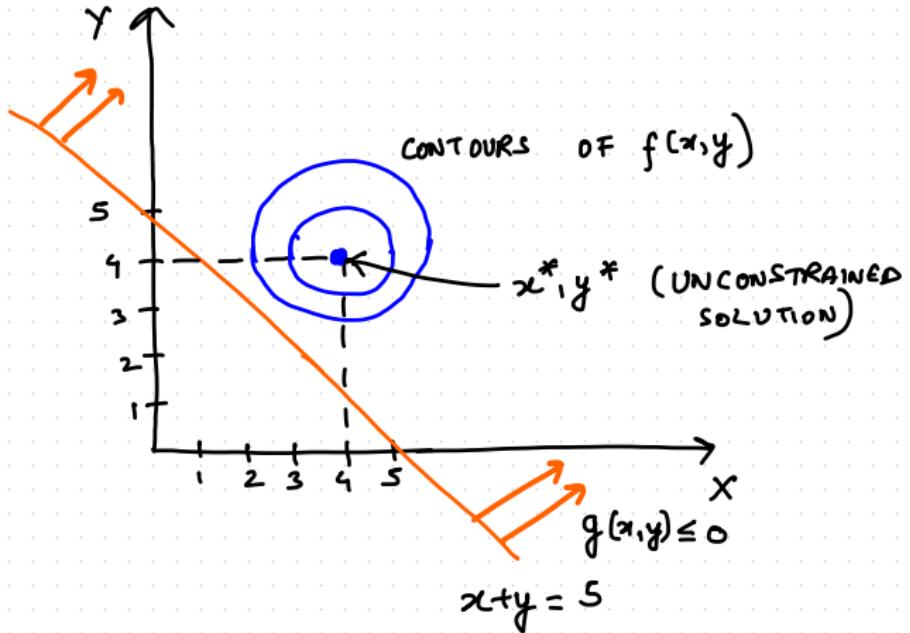
$$\text{Minimize } f(x, y) = (x-4)^2 + (y-4)^2$$
$$\text{s.t. } x+y \geq 5$$

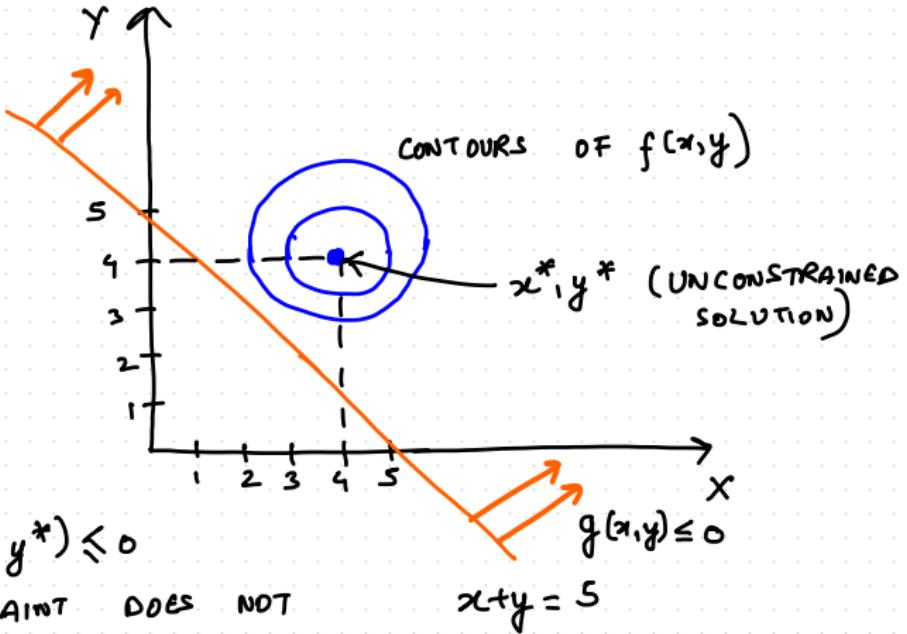
$$f(x, y) = (x-4)^2 + (y-4)^2$$
$$g(x, y) = -x - y + 5 \leq 0$$









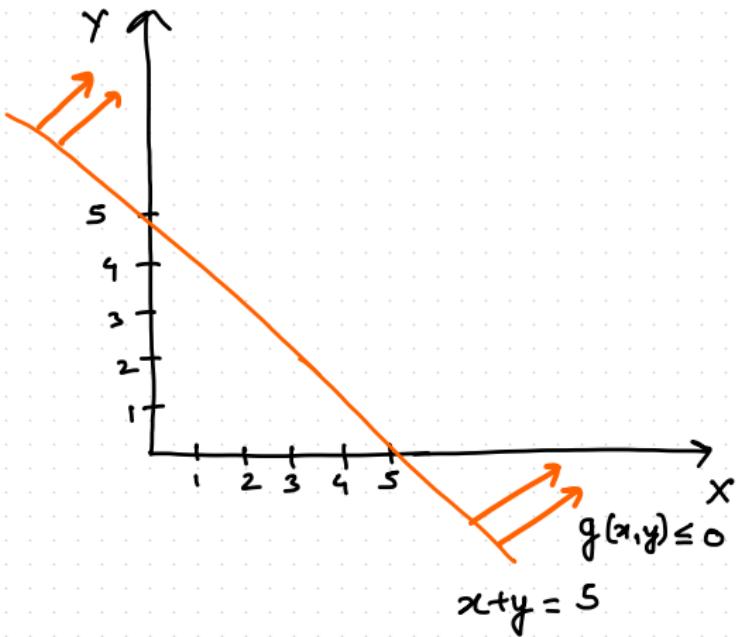


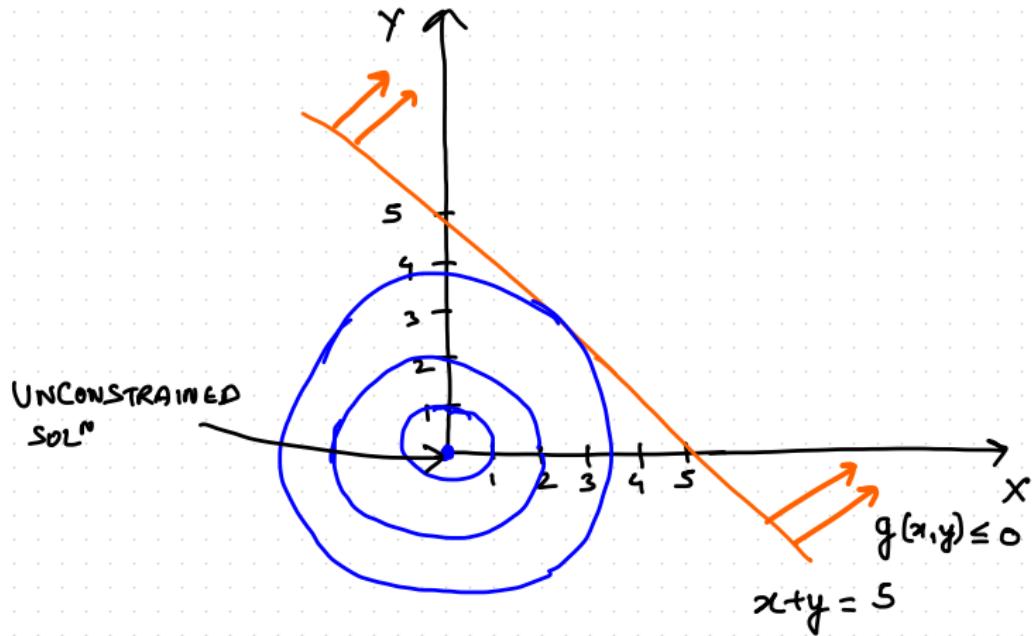
CONSTRAINT DOES NOT
TAKE PART

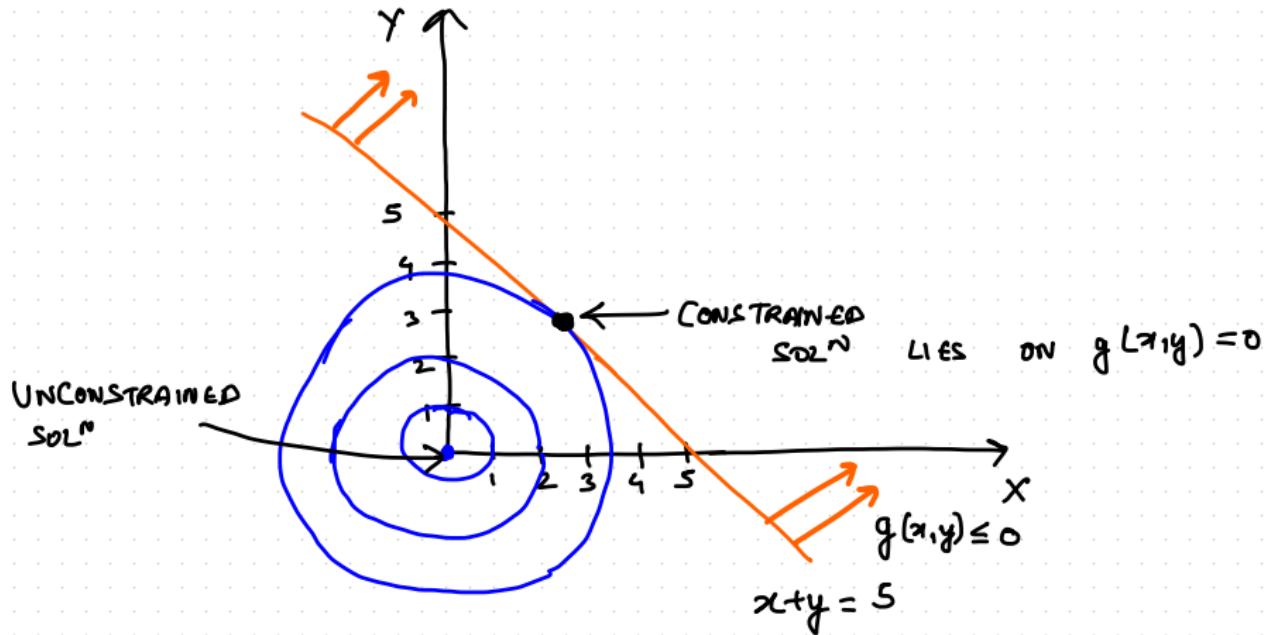
$$\Rightarrow \mu = 0 \Rightarrow \boxed{\mu g(x,y) = 0}$$

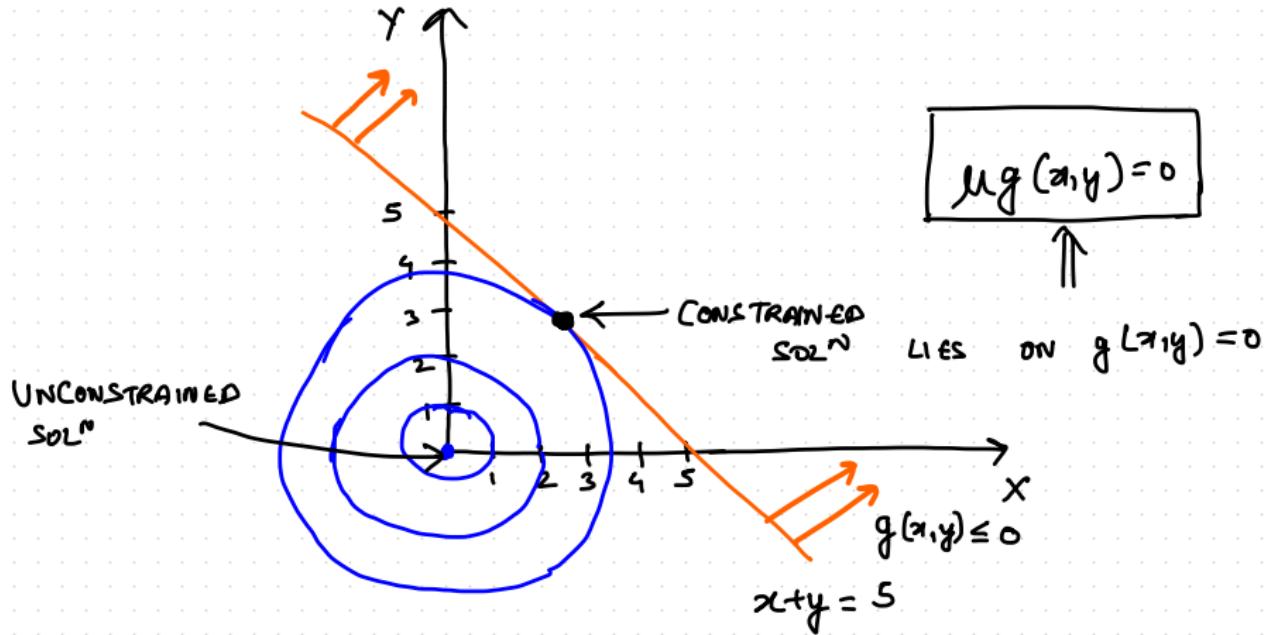
$$\text{Minimize } f(x, y) = x^2 + y^2$$
$$\text{s.t. } x+y \geq 5$$

$$f(x, y) = x^2 + y^2$$
$$g(x, y) = -x - y + 5 \leq 0$$



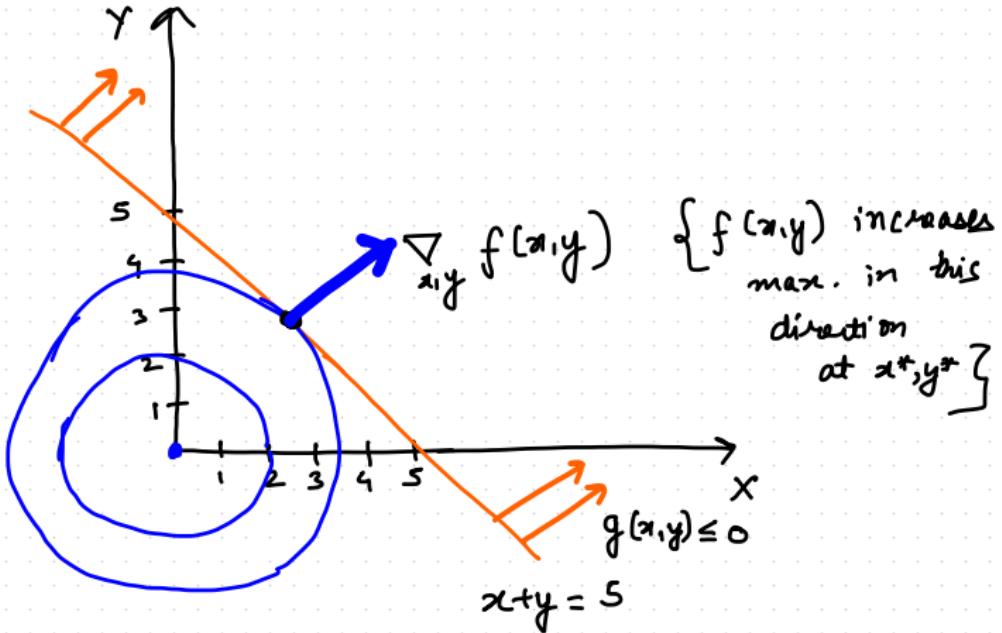


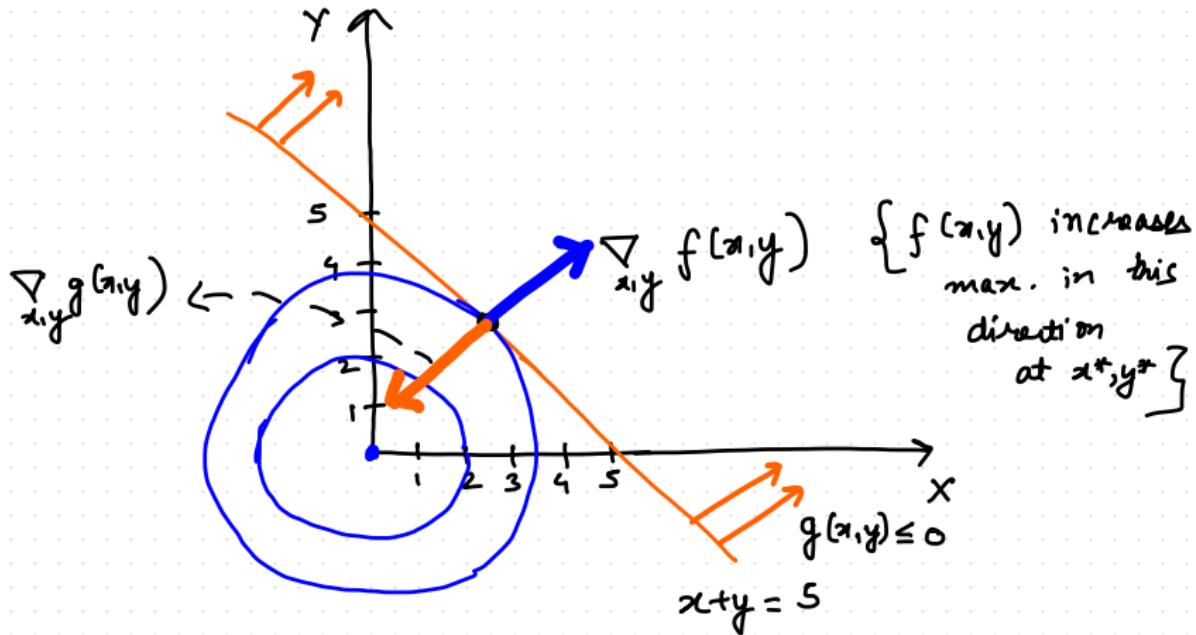


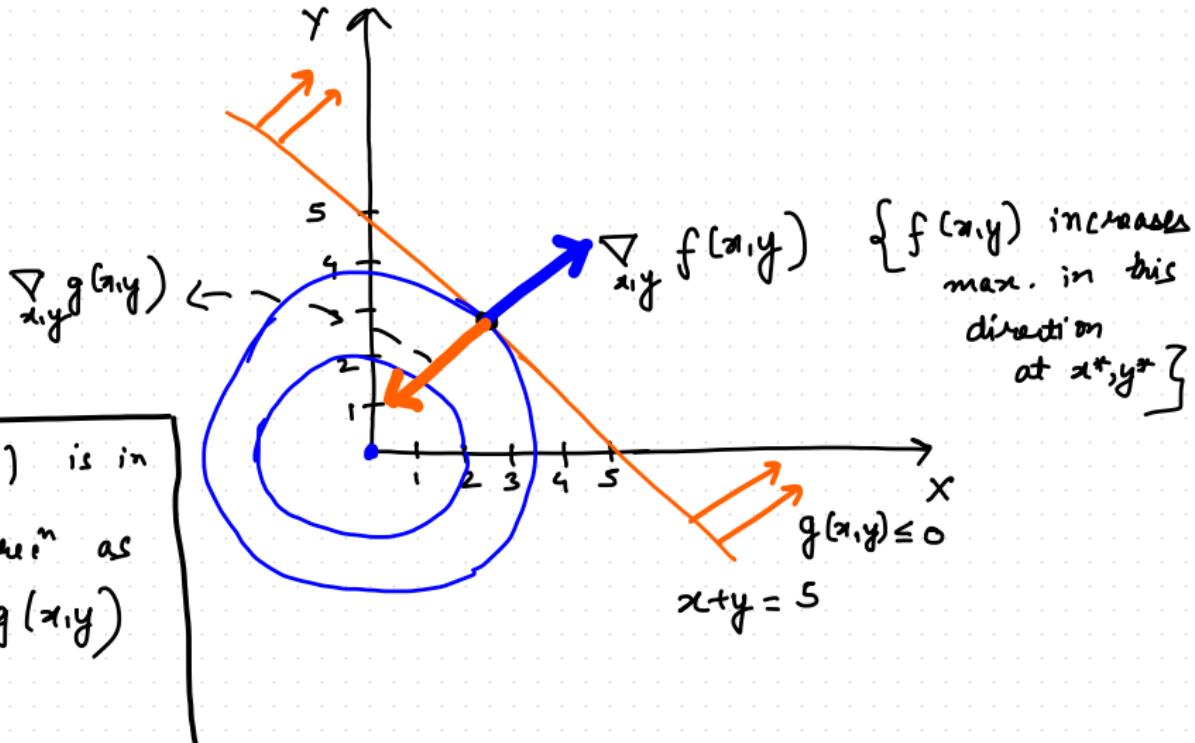


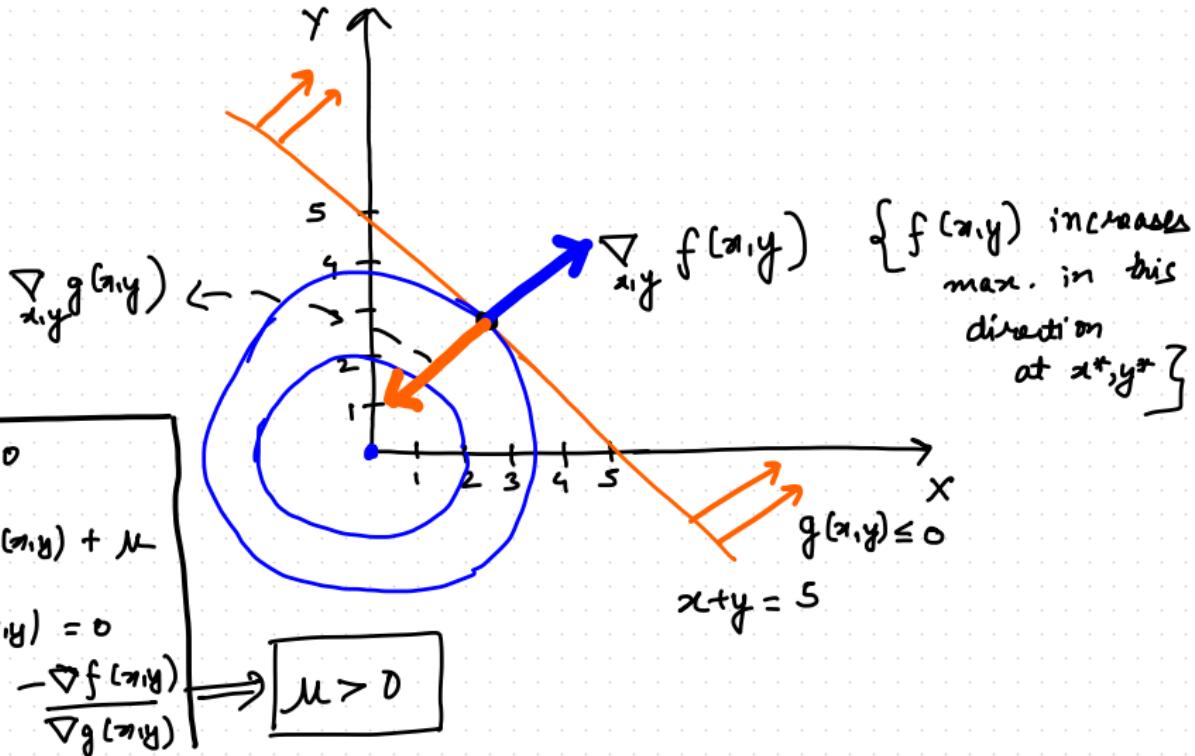
WHY $\mu_i > 0 \neq i$

CONSIDER CASE WHEN $\mu \neq 0$









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)

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

$$\mu_i \geq 0$$

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

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

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

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

$$g_2(x, y) = -x$$

Example

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

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

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

$$g_2(x, y) = -x$$

$$g_3(x, y) = -y$$

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

$$\nabla_x L(x, y, \lambda, \mu_1, \mu_2, \mu_3) = 0 \\ \implies 2x + \lambda + 2\mu_1 x - \mu_2 = 0 \dots \dots \dots \quad (1)$$

$$\begin{aligned} \nabla_y L(x, y, \lambda, \mu_1, \mu_2, \mu_3) &= 0 \\ \implies 2y + 2\lambda + 2\mu_1 y - \mu_3 &= 0 \end{aligned} \quad \dots \dots \dots \quad (2)$$

Example

Stationarity

$$\nabla_y L(x, y, \lambda, \mu_1, \mu_2, \mu_3) = 0 \\ \implies 2y + 2\lambda + 2\mu_1 y - \mu_3 = 0 \dots \dots \dots \quad (2)$$

Equality Constraint

Example

Stationarity

$$\begin{aligned}\nabla_x L(x, y, \lambda, \mu_1, \mu_2, \mu_3) &= 0 \\ \implies 2x + \lambda + 2\mu_1 x - \mu_2 &= 0 \dots \dots \dots \quad (1)\end{aligned}$$

$$\begin{aligned}\nabla_y L(x, y, \lambda, \mu_1, \mu_2, \mu_3) &= 0 \\ \implies 2y + 2\lambda + 2\mu_1 y - \mu_3 &= 0 \dots \dots \dots \quad (2)\end{aligned}$$

Equality Constraint

$$x + 2y = 4 \dots \dots \dots \quad (3)$$

Slackness

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

$$\mu_2 x = 0 \dots \dots \dots \quad (5)$$

$$\mu_3 y = 0 \dots \dots \dots \quad (6)$$

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,

$$x = 0.8$$

$$y = 1.6$$

$$\text{MINIMIZE} \quad x^2 + y^2$$

s.t.

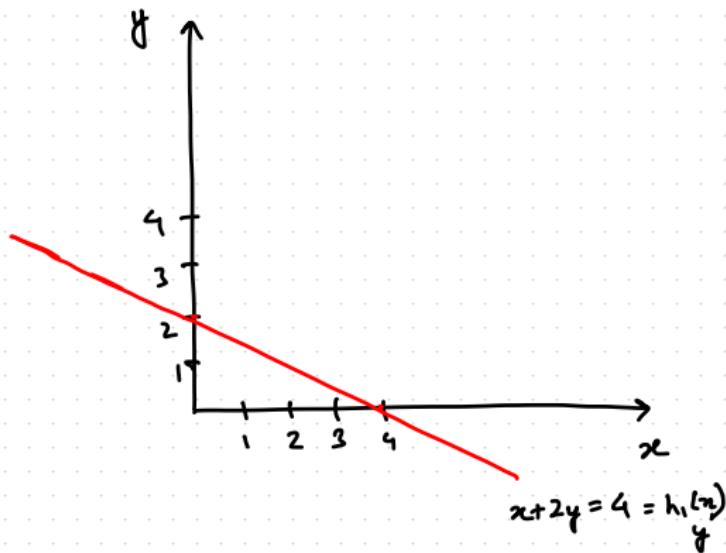
$$x^2 + y^2 \leq 5 \quad \text{or} \quad g_1(x,y) = x^2 + y^2 - 5 \leq 0 \quad (\mu_1)$$

$$x + 2y = 4 \quad \text{or} \quad h(x,y) = x + 2y - 4 = 0$$

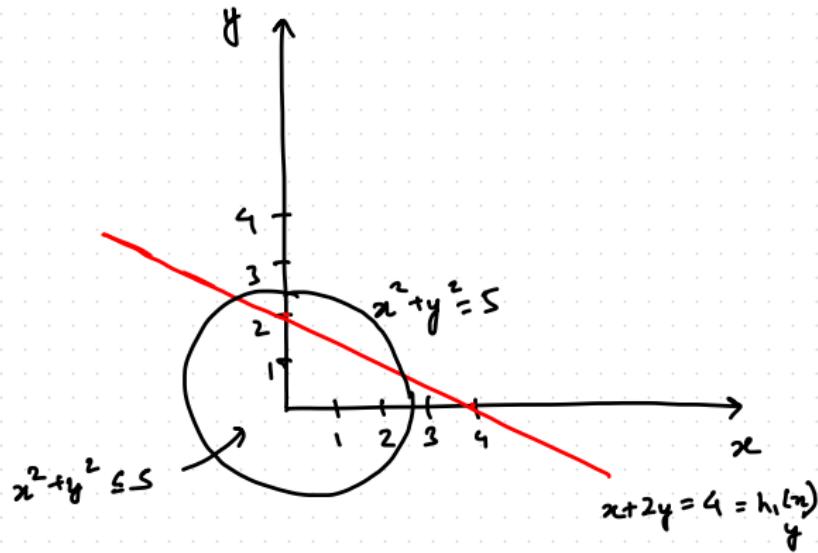
$$x, y \geq 0 \quad \text{or} \quad g_2(x,y) = -x \leq 0 \quad (\mu_2)$$

$$g_3(x,y) = -y \leq 0 \quad (\text{multiplier: } \mu_3)$$

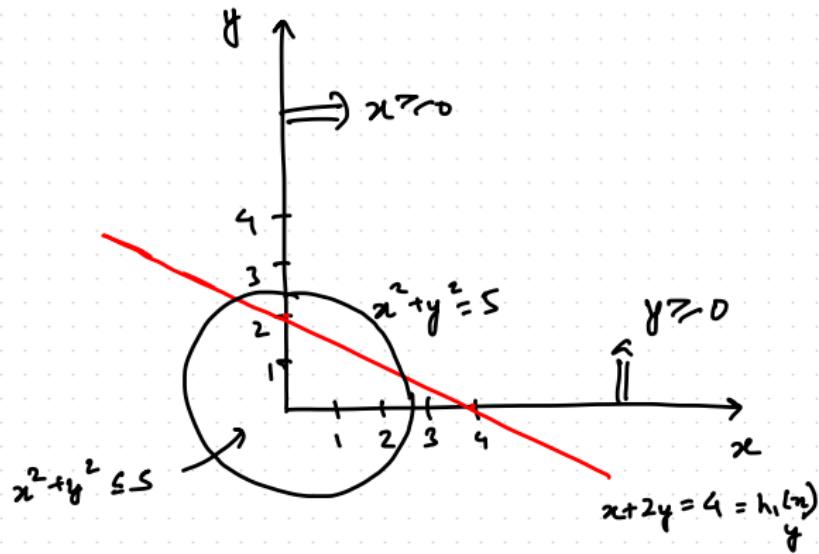
MINIMIZE $x^2 + y^2$



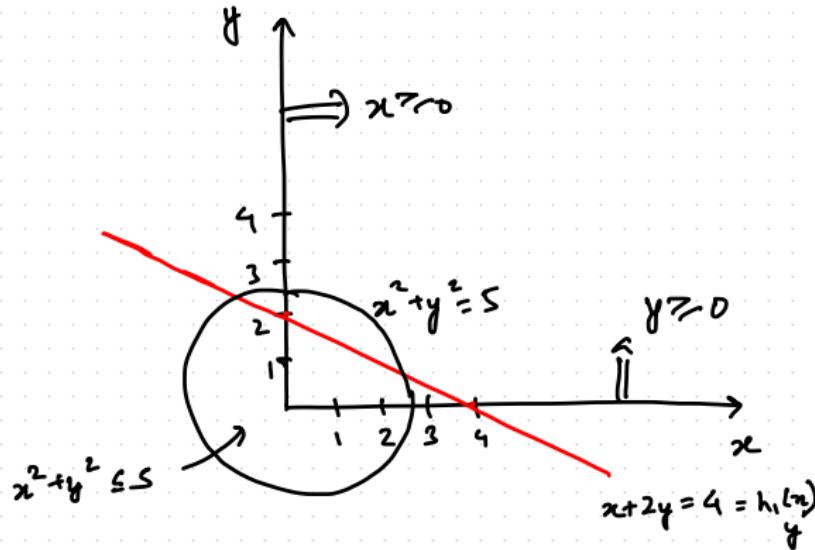
MINIMIZE $x^2 + y^2$



MINIMIZE $x^2 + y^2$

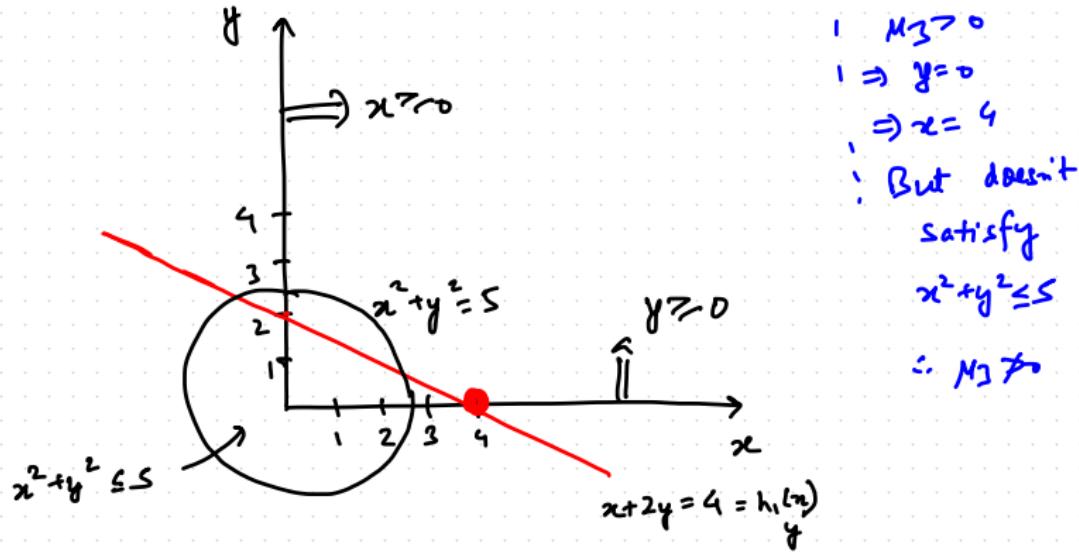


MINIMIZE $x^2 + y^2$



$\mu_3 = 0$ $\lambda > 0$

MINIMIZE $x^2 + y^2$



$M_3 = 0$ $\because \alpha > 0$

$M_3 > 0$

$\Rightarrow y = 0$

$\Rightarrow x = 4$

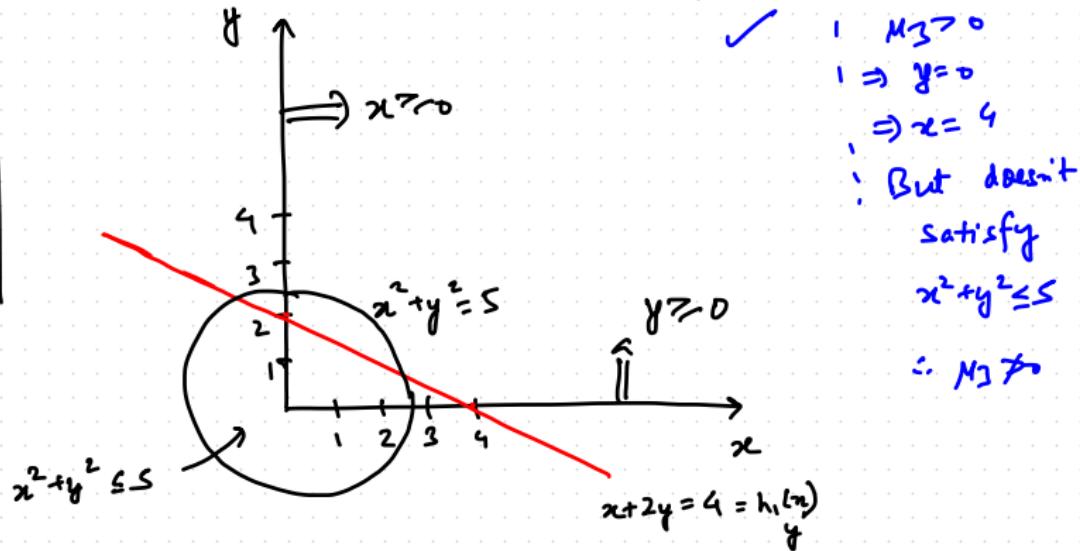
\therefore But doesn't satisfy

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

$\therefore M_3 \neq 0$

MINIMIZE $x^2 + y^2$

$\mu_3 = 0$



$\mu_3 = 0$, $\alpha > 0$

$\checkmark \quad \mu_3 > 0$

$\Rightarrow y = 0$

$\Rightarrow x = 4$

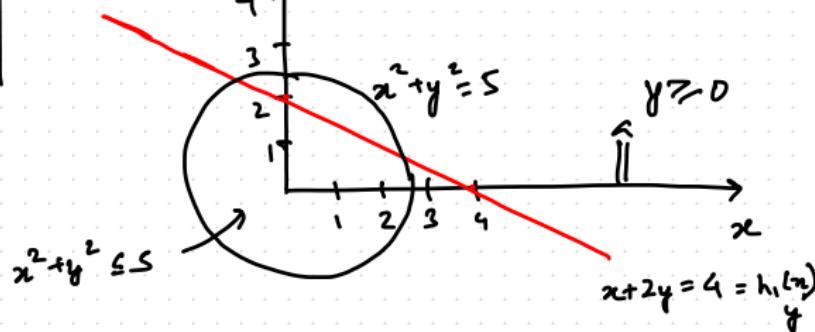
But doesn't satisfy

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

$\therefore \mu_3 \neq 0$

MINIMIZE $x^2 + y^2$

$\mu_3 = 0$

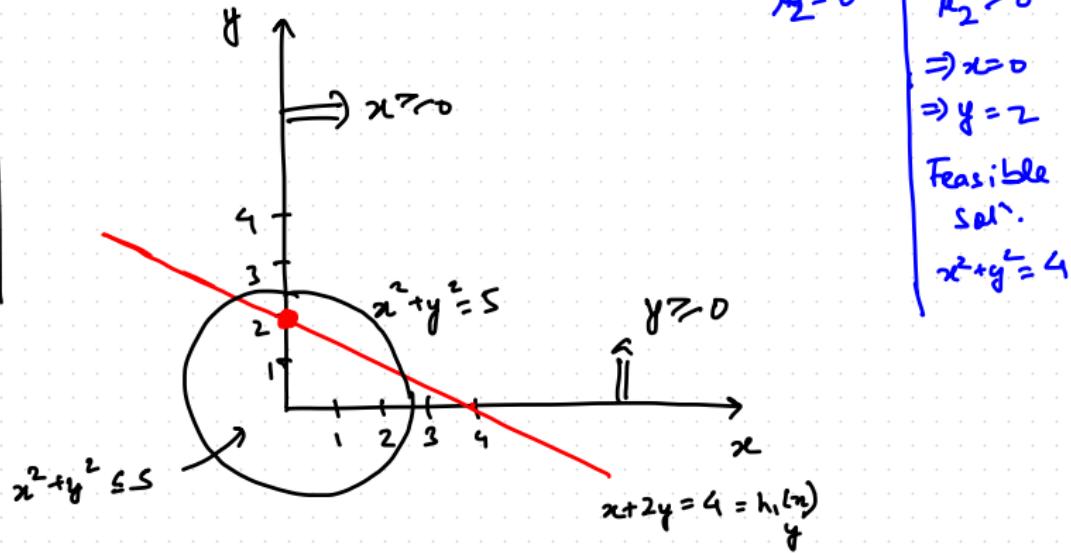


$\mu_2 = 0$

$\mu_2 > 0$

$$\text{MINIMIZE} \quad x^2 + y^2$$

$$\mu_3 = 0$$



$$\mu_2 = 0$$

$$\mu_2 > 0$$

$$\Rightarrow x = 0$$

$$\Rightarrow y = 2$$

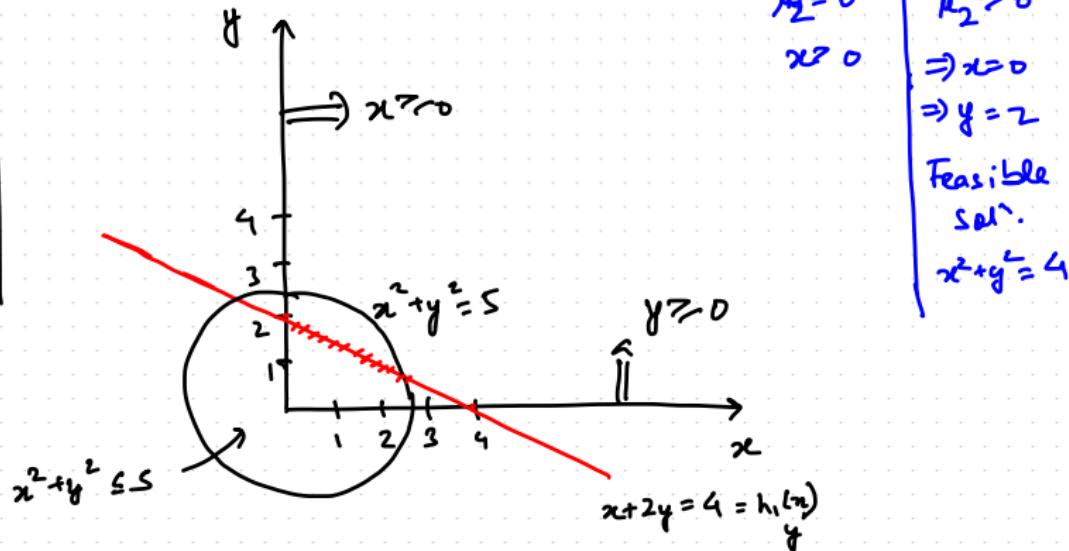
Feasible

Solⁿ.

$$x^2 + y^2 = 4$$

MINIMIZE $x^2 + y^2$

$\mu_3 = 0$

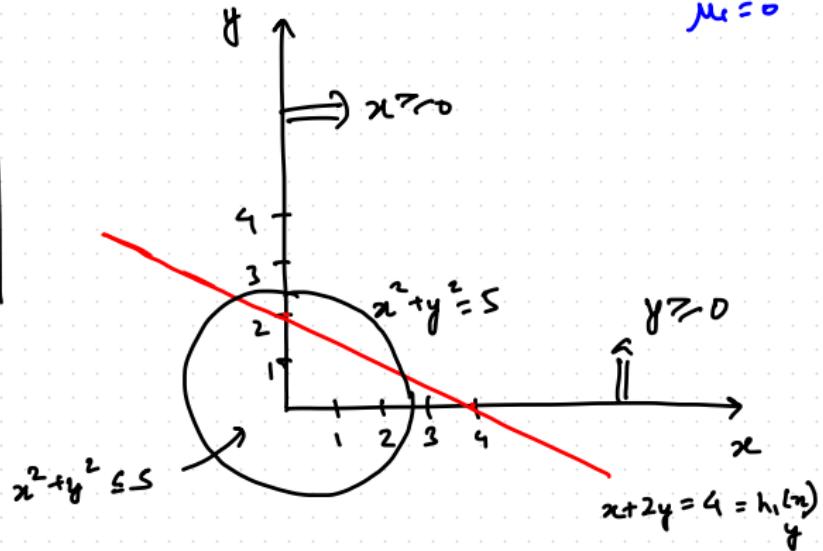


$\mu_2 = 0$
 $x \geq 0$

$\mu_2 > 0$
 $\Rightarrow x = 0$
 $\Rightarrow y = 2$
 Feasible
 Sol.
 $x^2 + y^2 = 4$

MINIMIZE $x^2 + y^2$

$$\mu_3 = 0$$



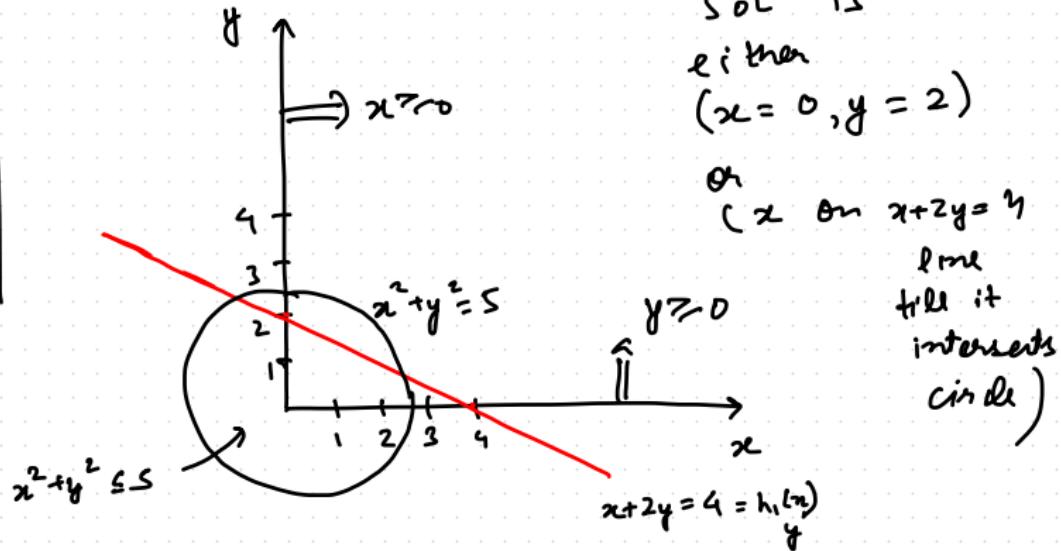
$$\mu_1 = 0$$

$$\begin{aligned} \mu_1 &> 0 \\ \Rightarrow x^2 + y^2 &= 5 \end{aligned}$$

But we have
seen
better
 μ_1^2
 $\therefore \mu_1 = 0$

MINIMIZE $x^2 + y^2$

$\mu_3 = 0$
 $\mu_4 = 0$

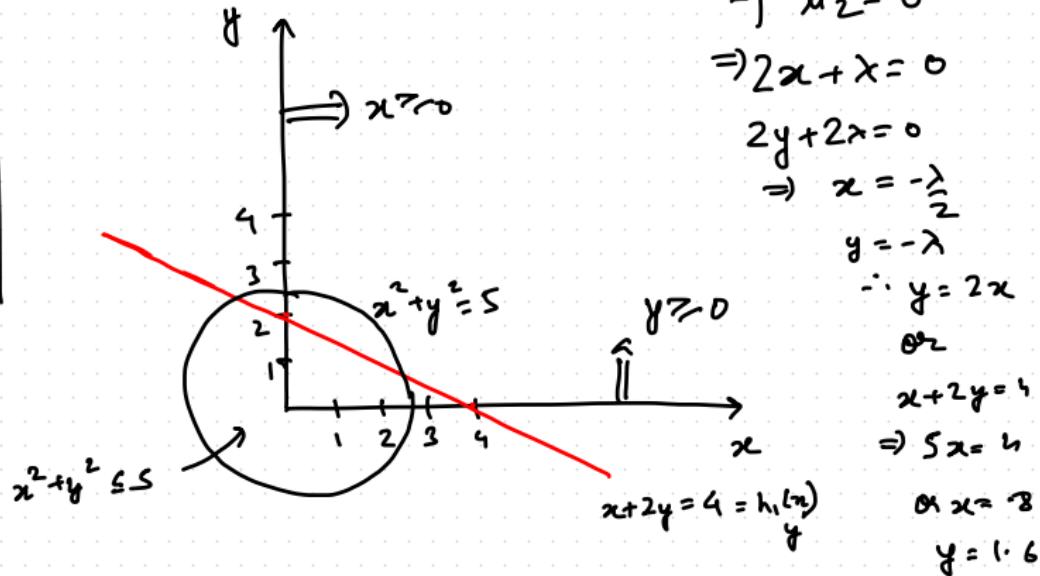


SOLⁿ is
 either
 $(x = 0, y = 2)$
 or
 $(x \text{ on } x+2y=4$
 line
 till it
 intersects
 circle)

$$x + 2y = 4 = h_1(x, y)$$

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
 $x + 2y = 4$
 $\Rightarrow 5x = 4$
 $\text{or } x = 0.8$
 $y = 1.6$

