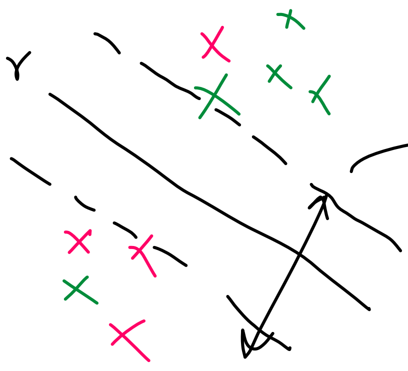
Problem with SVC

14 July 2023 09:15

✓ Maximal margin classifier



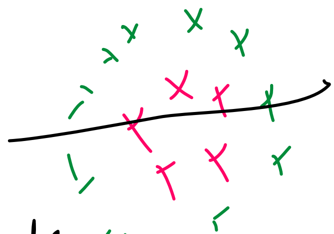
→ Hard margin SVM

↓

✓ SVC { Soft margin SVM } → linear dataset

↓

$\boxed{\text{SVM}}$ → SVC + Kernels → non linear



↓

Kernels

Kernels Intuition

14 July 2023 09:15

age	side effect
72	1
32	0
26	0
5	1

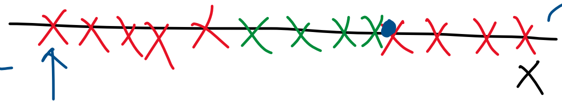
$\boxed{SVC}$ ↓

2D – line
3D → hyper

1D – point

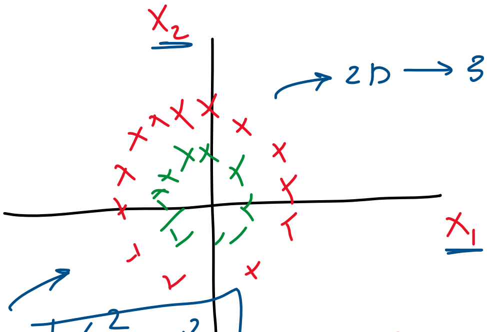
x

1D → 2D



age

$\underline{x_1^2 + x_2^2}$ ↑



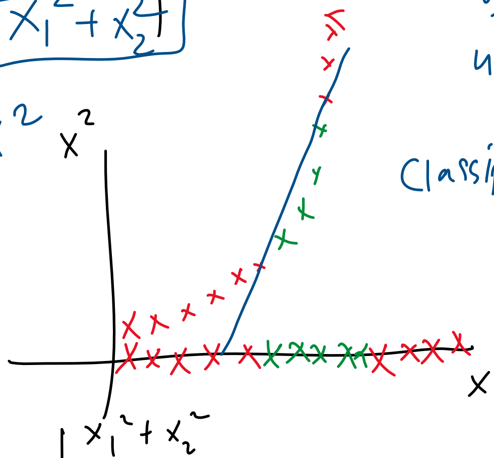
2D → 3D

1) input → <u>higher dim</u>
   <u>kernels</u>
2) linearly separable
3) SVC
4) project down

functions
<u>math</u>

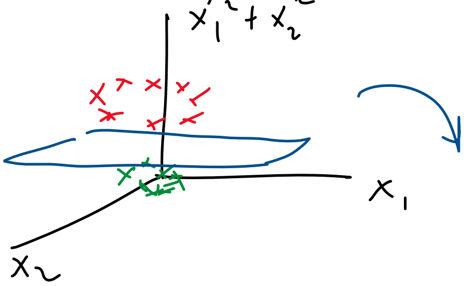
$\boxed{x_1^2 + x_2^2}$

kernel → x^2 x^2



Classify

$\underline{e^{-x^2}}$ $e^{(x_1^2+x_2^2)}$



Code

14 July 2023 10:31

Types of Kernels

14 July 2023 10:31

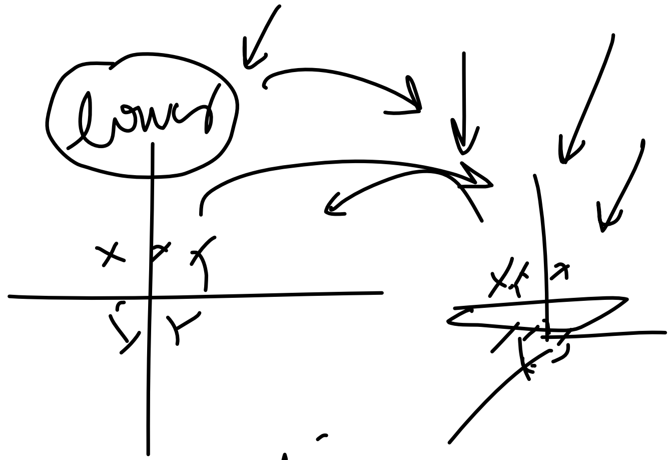
- linear
- rbf
- poly
- sigmoid

Why is it called Trick?

14 July 2023 10:32

Kernel $\boxed{\text{trick}}$?

fast



computation

The gradient of a function at a point is a vector that points in the direction of the steepest ascent of the function at that point. The magnitude (or length) of the gradient vector is equal to the rate of increase of the function in that direction.

Contour lines (or level sets) of a function are curves that connect points where the function has the same value. For a 2D function, these are like the lines of constant altitude on a topographic map.

There are a few key relationships between gradients and contour lines:

1. The gradient at a point is perpendicular (or orthogonal) to the contour line passing through that point. This is because the contour line represents the direction of no change in the function value, while the gradient represents the direction of maximum change.
2. The gradient points in the direction where the function increases most rapidly. If you were to walk along the contour line (where the function value doesn't change), the direction you'd need to go to start climbing as steeply as possible is the direction of the gradient.
3. The magnitude of the gradient (how long the gradient vector is) indicates how steeply the function is increasing. If the contour lines are close together, that means the function is changing rapidly, so the gradient is large. If the contour lines are far apart, the function is changing slowly, so the gradient is small.

Mathematics of SVM

14 July 2023 20:27

optimization prob

constrained optimization problem

linear reg → OLS, GD

Logistic Reg

minimize

$$L = \underset{\beta}{\arg\min} \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

min

$$\frac{\partial L}{\partial \beta} = 0$$

constraint

min

$$\underset{A,B,C}{\arg\min} \frac{\sqrt{A^2+B^2}}{2} + C\frac{1}{n}\sum_{i=1}^{n}\xi_i$$

n

such that

$$y_i(Ax_{1i} + Bx_{2i} + C) \geq 1 - \xi_i$$

← constrained optimization

↳ restriction

→ main function

$$\underset{x,y}{\arg\max}\ \boxed{x^2y}$$

such that

$$\boxed{x^2+y^2=1}$$ → constraint

x, y

$\boxed{x = 5 \quad y = 7}$ ×

$x = \frac{1}{\sqrt{2}} \quad y = \frac{1}{\sqrt{2}}$

$(5)^2 \times 7 = 175$

$\frac{1}{2} \times \frac{1}{\sqrt{2}} = \boxed{\frac{1}{2\sqrt{2}}}$

$(5)^2 + (7)^2 = 1$

$\frac{1}{2} + \frac{1}{2} = 1$

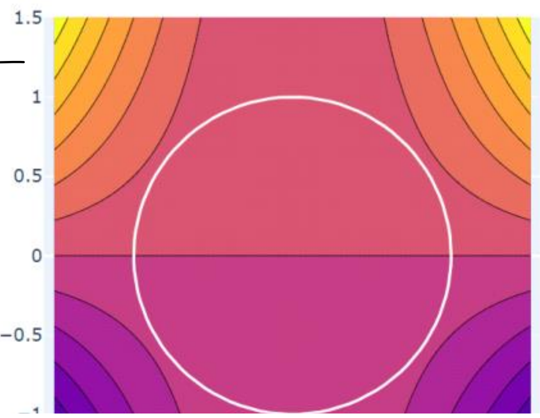
$$\underset{x,y}{\arg\max}\ \boxed{x^2y}$$

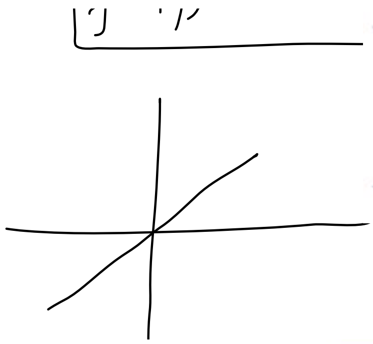
such that

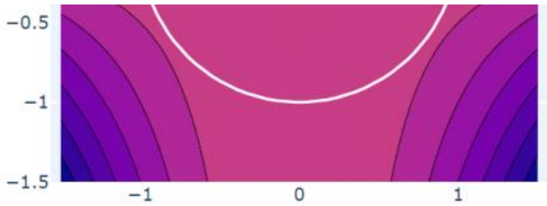
$$x^2+y^2 =$$

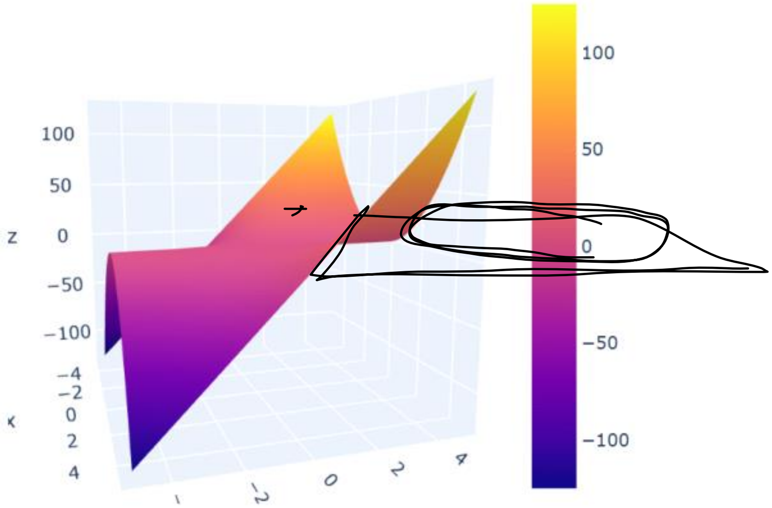
geometric intuition

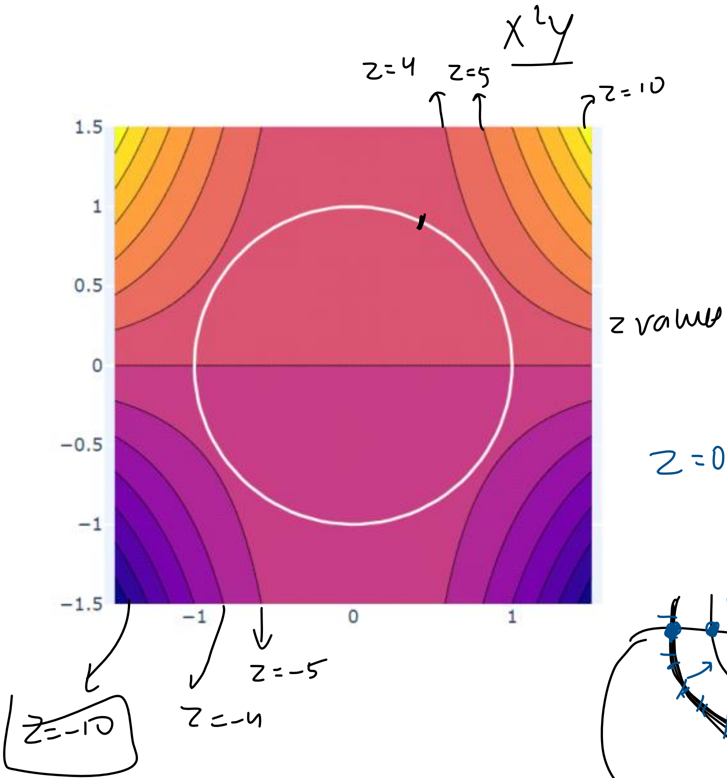
$$f(x,y) = z = x$$



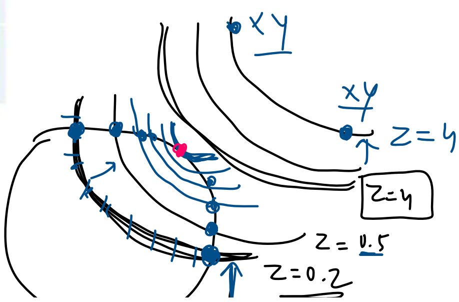






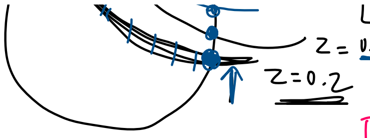


z=0.2



x,y
acceptable

$\boxed{z=-10}$ $z=-4$

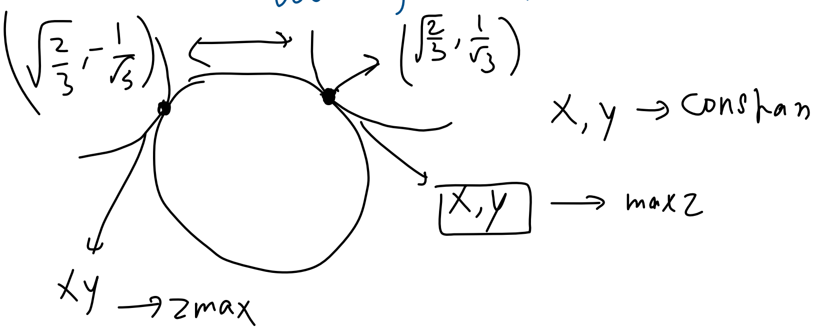


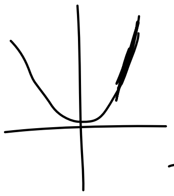
$\boxed{x, y}$ acceptable

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

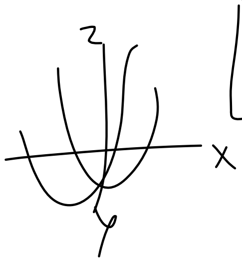
constraint : $\boxed{x^2+y^2=1}$ ~~x~~ true

$z = 0.2 \rightarrow$ highest possible value of $z\ (x,y)$





$\frac{dy}{dx}$

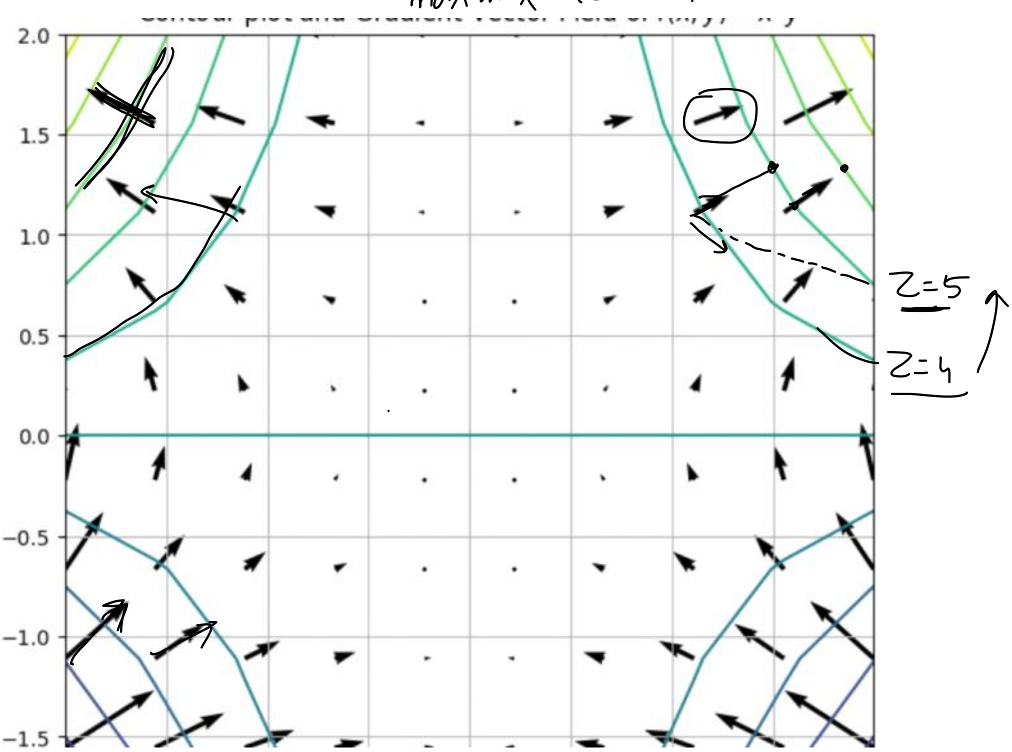


$\begin{bmatrix} \frac{\partial z}{\partial x} & \frac{\partial z}{\partial y} \end{bmatrix}$

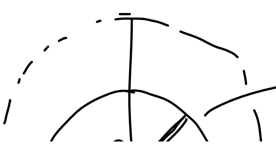
↳ gradient

direction of maximum ascent / change

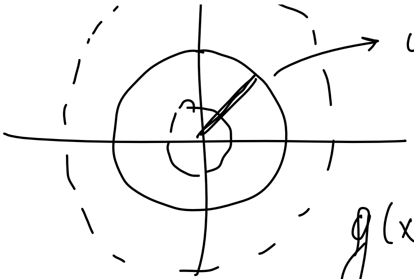
perpendicular



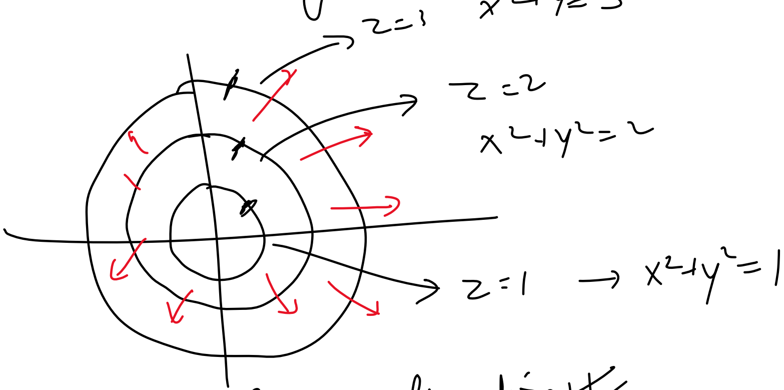
$x^2+y^2=1$ $\quad x^2+y^2=2 \quad x^2+y^2=0.5$

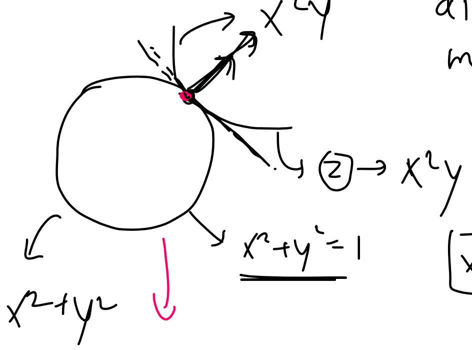


circle = 1



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





direction ✓

magn ✗

$\boxed{x, y}$ $\boxed{\max z}$

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

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

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

λ max

$$\boxed{\nabla f(x,y) = \lambda \nabla g(x,y)} \rightarrow$$

λ → scalar

→ lagrange's multiplier

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

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

$$\rightarrow \begin{bmatrix} 2xy \\ \end{bmatrix}$$

$x^2y \rightarrow \frac{\partial f}{\partial x} = 2xy$

$\frac{\partial f}{\partial y} = x^2$

$x^2+y^2 \rightarrow \frac{\partial g}{\partial x} = 2x$

$$\hookrightarrow \begin{bmatrix} 2xy \\ x^2 \end{bmatrix}$$

$$x^2+y^2 \rightarrow \frac{\partial g}{\partial x} = 2x$$

$$\frac{\partial g}{\partial y} = 2y$$

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

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

$\boxed{\cancel{2xy = \lambda 2x}}$ — 1

$\boxed{x^2 = \lambda 2y}$ — 2

$x, y, \cancel{z}, \lambda$

$x^2+y^2=1$ — 3

$\boxed{y=\lambda}$

$\boxed{x^2 = 2y^2}$

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

$3y^2=1 \qquad y^2 = \pm\frac{1}{\sqrt{3}}$

$x^2+\frac{1}{3}=1$

$x^2 = 1-\frac{1}{3} \qquad x^2=\frac{2}{3} \qquad x = \pm\sqrt{\frac{2}{3}}$

$$x = \sqrt{\frac{2}{3}}, -\sqrt{\frac{2}{3}}$$

$Z \rightarrow$ maximum

$$y = \frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}$$

$x \quad y$

$\boxed{\left(\sqrt{\frac{2}{3}}, \frac{1}{\sqrt{3}}\right)}$ $\left(\sqrt{\frac{2}{3}}, -\frac{1}{\sqrt{3}}\right)$ X $\boxed{\left(-\sqrt{\frac{2}{3}}, \frac{1}{\sqrt{3}}\right)}$ $\left(-\sqrt{\frac{2}{3}}, -\frac{1}{\sqrt{3}}\right)$ X

$Z = x^2 y \rightarrow \frac{2}{3}\frac{1}{\sqrt{3}} = \frac{2}{3\sqrt{3}}$

$\frac{2}{3}\frac{1}{\sqrt{3}} = \frac{2}{3\sqrt{3}}$

$x, y \rightarrow z$ max

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

$\left(\sqrt{\frac{2}{3}}\right)^2 + \left(\frac{1}{\sqrt{3}}\right)^2$

$\frac{2}{3} + \frac{1}{3} = \frac{3}{3} = 1$

Lagrangian multiplier $\rightarrow$ SVM

$x^2 y$ $\quad$ $x^2 + y^2$

$\left[\underset{x,y}{\text{argmax}}\ x^2 y \quad ST \quad x^2 + y^2 = 1\right] \rightarrow \boxed{\nabla f(x,y) = \lambda g(x,y)}$

optimazation prob

$x^2 y$ $\quad$ $y^2 + y^2$

$\left[L(x,y,\lambda) = \underset{x,y}{\text{argmax}}\ f(x,y) - \lambda\,(g(x,y) - 1)\right]$

$\frac{\partial L}{\partial x} = 0$ $\quad$ $\boxed{\frac{\partial L}{\partial y}} = 0$ $\quad$ $\boxed{\frac{\partial L}{\partial \lambda}} = 0$ $\quad$ $\frac{\partial f}{\partial x} - \lambda \frac{\partial g}{\partial x} = 0$

1 eq. $\quad$ 2nd $\quad$ 3rd

$\boxed{\frac{\partial f}{\partial x} = \lambda \frac{\partial g}{\partial x}}$

$\boxed{2xy = \lambda 2x}$

$g(x,y) - 1 = 0$

$g(x,y) = 1$

$\boxed{x^2 + y^2 = 1}$

$\frac{\partial f}{\partial y} - \lambda \frac{\partial g}{\partial y} = 0$

$x^2 - \lambda 2y = 0$

$\boxed{x^2 = \lambda 2y}$

$\sum_{i=1} (x_i - y_i)$

$(x_1 - y_1) + (x_2 - y_2)$

$\left[\underset{x,y}{\text{argmax}}\ x^2 y \quad ST \quad x^2 + y^2 = 1\right]$ const opt

$\rightarrow$ computer

optim-

Optim-

$$\boxed{L(x,y,\lambda) = \underset{x,y,\lambda}{\arg\max} \; x^2 y - \lambda(x^2+y^2-1)}$$

largarge multip

SVM dual problem