Consider $Ax = b$ where no Solu exits.
Inconsistent

$b \notin \text{colsp}(A)$.

$\|b - A\hat{x}\|^2$ must be minimum.

$\Rightarrow (b - A\hat{x})^T (b - A\hat{x})$

$= (b^T - (A\hat{x})^T)(b - A\hat{x})$.

$\frac{df}{d\vec{x}} = \frac{d}{d\vec{x}}\left(b^T b - b^T A\hat{x} - (A\hat{x})^T b + (A\hat{x})^T A\hat{x}\right)$

$= \frac{d}{d\vec{x}}\left(b^T b - 2b^T A\hat{x} + x^T A^T A\, x.\right).$

$= \quad 0 \quad - 2A^T b + 2(A^T A)x.$

$= \quad 0 = -2A^T b + \cancel{2}(A^T A)\, x = 0.$

$1 \times n.$

$y = f(x)$

$\frac{dy}{dx} = 0$

$\frac{d^2y}{dx^2}$ at $\frac{dy}{dx} = 0$

If $> 0 \Rightarrow$ Min

$<$ Max.

$x^T y = y^T x.$

$b^T A x = (Ax)^T b$

$=$

$x^T A x$

$A^T A = L$

$\frac{d}{dx} x^T L x.$

$= \underbrace{(L + L^T)}_{2Lx.} x$

$A^{m \times n}$

$A^T = n \times m$

$= (A^T A)\hat{x} = A^T b$ $n \times 1$

$$\boxed{\hat{x} = \boxed{(A^T A)^{-1} A^T b}}$$

$x \rightarrow \underset{x \in \mathbb{R}^n}{\text{Min}} \; \underline{\|Ax - b\|^2}.$

$x = (A^T A)^{-1} A^T b.$ Minimizer sq.

$Ax = b$ $\quad A\hat{x} = A(A^T A)^{-1} A^T b.$

Projection Matrix

$$\begin{bmatrix} b_1 & b_2 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$b_1(a_{11}x_1 + a_{12}x_2) + b_2(a_{21}x_1 + a_{22}x_2)$$

$$\frac{\partial}{\partial x_1} = \underbrace{b_{11}a_{11} + b_2 a_{21}}$$

$$\frac{\partial}{\partial x_2} = b_1 a_{12} + b_2 a_{22}.$$

$$\nabla = \begin{bmatrix} \downarrow \\ , \end{bmatrix} \cdot (\;)$$

$$= \begin{bmatrix} b_1 a_{11} + b_2 a_{21} \\ b_1 a_{12} + b_2 a_{22} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

$$A^T b$$

Lagrangian Multipliers. → Converts a constrained optimizn to unconstrained.

~~$Ax = \lambda x$~~

~~$A^T x = \lambda x$~~

~~A : RSM.~~

~~$A^T = A$.~~

~~x : Unit vector~~

~~$A^T A x = A^T(\lambda x) = \lambda(A^T x) = \lambda(Ax) = \lambda^2 x$.~~

~~$x^T(A^T A)x = x^T \lambda^2 x$.~~

~~$= \lambda^2 x^T x$~~

~~$= \lambda^2$.~~

$A = PDP^{-1}$

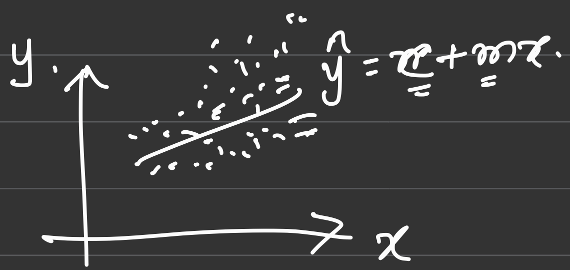
for a Cov Matrix $D = P^{-1}AP$.

Optimization:

Objective fn $f(x)$ Cost function.

Minimize or Maximize. $f(\underline{x})$

Cost fn: $f(\vec{x}) \rightarrow \underset{\vec{x}}{\text{Min}}$ $\quad \vec{x} \in \mathbb{R}^n$.



Linear Regression $\underline{\underline{y}} = c + mx$.

Original y values $\hat{y} = c + mx$

$$\theta_1 + \theta_2 x_1 + \theta_3 x_2 + \dots + \theta_n x_n$$

for $x \in \mathbb{R}^n$, $\hat{y} = \theta_1 f_1(\vec{x}) + \theta_2 f_2(\vec{x}) + \dots + \theta_n f_n(\vec{x})$.

Feature Vector

value.

$$f_i(\vec{x}) = f: \mathbb{R}^n \longmapsto \mathbb{R}.$$

$\{x_1, \dots, x_n\}$

$A^{m\times n} x^{n\times 1} = b^{m\times 1}$. $\quad x_1 \dots x_n$ be variables

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}. \qquad \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \vdots \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ \vdots \end{bmatrix}.$$

$f_1(\vec{x}) = a x_1 + b x_2 \qquad x \in \mathbb{R}^2$

$\| \vec{y} - \vec{\hat{y}} \|^2 \rightarrow$ Minimize.

x	y
-1	1
1	3
2	0

$(A^T A)^{-1} A^T b.$

$y = c + m x$

$$\begin{cases} y_1 = c + m x_1 \\ y_2 = c + m x_2 \\ y_3 = c + m x_3 \end{cases}$$

$$\underset{b}{\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}} = \underset{A}{\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \end{bmatrix}} \underset{x}{\begin{bmatrix} c \\ m \end{bmatrix}}$$

$$\begin{cases} \hat{y}_1 = \hat{c} + \hat{m} x_1 \\ \hat{y}_2 = \hat{c} + \hat{m} x_2 \\ \hat{y}_3 = c + m x_3 \end{cases}$$

$$\boxed{x^{+} = (A^T A)^{-1} A^T b}$$

$\hat{y} = \hat{c} + \hat{m} x \rightarrow$ Model.

$$\vec{r} = \begin{bmatrix} y_1 - \hat{y}_1 \\ y_2 - \hat{y}_2 \\ y_3 - \hat{y}_3 \end{bmatrix}$$

$\|\vec{r}\|^2$ very small

$\rightarrow$ Least.

Max Marks in MAAI $(\underline{x_1} + \underline{x_2} + \cdots + \underline{x_5})$

Subject to prepn time ≤ 24 hrs.

* $y = f(x)$ $\quad \frac{dy}{dx} = 0$ $\quad \frac{d^2y}{dx^2}$ at that $\underline{x}$ where $\frac{dy}{dx} = 0$.

* Let $z = f(x_1, x_2)$ which is differentiable at (x_0, y_0).

Critical point: (x_0, y_0) is a critical point of f if

either

a) $\frac{\partial f}{\partial x_1}(x_0, y_0) = \frac{\partial f}{\partial x_2}(x_0, y_0) = 0$

or b) the partial derivatives $\frac{\partial f}{\partial x_1}$ or $\frac{\partial f}{\partial x_2}$ does not exist.

SADDLE Point is a point on a graph where the slopes are all 0, but the point at which the slope is 0 is neither a local max nor a min.

Q: $f(x, y) = x^3 + 12x^2 + 6y^2 - 24xy + 18x.$

Critical points of f: $\frac{\partial f}{\partial x} = 0$; $\frac{\partial f}{\partial y} = 0.$

$$\frac{\partial f}{\partial x} = 3x^2 + 24x - 24y + 18 = 0 \qquad \frac{\partial f}{\partial y} = 12y - 24x = 0$$

$$y - 2x = 0$$

$$\underline{\underline{y}} = 2x.$$

$$3x^2 + 24x - 48x + 18 = 0.$$

$$3x^2 - 24x + 18 = 0$$

$$x^2 - 8x + 6 = 0.$$

$$x = \frac{8 \pm \sqrt{64 - 24}}{2} = \frac{8 \pm \sqrt{40}}{2} = \underline{\underline{4 \pm \sqrt{10}}}.$$

$$H: \begin{bmatrix} \frac{\partial}{\partial x}\left(\frac{\partial}{\partial x}\right) & \frac{\partial}{\partial x}\left(\frac{\partial}{\partial y}\right) \\ \frac{\partial}{\partial y}\left(\frac{\partial}{\partial x}\right) & \frac{\partial}{\partial y}\left(\frac{\partial}{\partial y}\right) \end{bmatrix} = \begin{bmatrix} 6x+24 & -24 \\ -24 & 12 \end{bmatrix}$$

Critical points: $(x, 2x)$.

$$= \left(4+\sqrt{10},\ 8+2\sqrt{10}\right) \qquad \left(4-\sqrt{10}\ \ 8-2\sqrt{10}\right)$$

$$H: \begin{bmatrix} 6(4+\sqrt{10}) & -24 \\ -24 & 12 \end{bmatrix} \qquad \begin{bmatrix} 6(4-\sqrt{10}) & -24 \\ -24 & 12 \end{bmatrix}$$

$$= \begin{bmatrix} 24+6(3.1) & -24 \\ -24 & 12 \end{bmatrix} \qquad \begin{bmatrix} 24-6(3.1) & -24 \\ -24 & 12 \end{bmatrix}$$

For minima Hessian must be positive definite
Maxima " " " Negative definite

$$\begin{bmatrix} 42.6 & -24 \\ -24 & 12 \end{bmatrix}$$

$\Downarrow$

det negative.

Saddle points.

$$\begin{bmatrix} 5.4 & -24 \\ -24 & 12 \end{bmatrix}$$

$64.8 - 576$

det < 0

$\Rightarrow$ One of the eigen val is neg & other is positive

Indefinite.

eigenval are opposite signs.

$\Rightarrow$ SADDLE POINT.

$f(x_1, x_2, x_3) = 2x_1^2 + 2x_2^2 + x_3^2 - x_1 - x_2 - 2x_1 x_2.$

Find the stationary points. Find if the stationary pt is max or min.

$\frac{\partial}{\partial x_1} = 4x_1 - 2x_2 - 1$ $\qquad$ $\frac{\partial}{\partial x_2} = 4x_2 - 2x_1 - 1$ $\qquad$ $\frac{\partial}{\partial x_3} = 2x_3$

$= 0$ $\qquad$ $= 0$ $\qquad$ $= 0$

$\Rightarrow x_3 = 0.$

$4x_1 - 2x_2 - 1 = 0$

$-2x_1 + 4x_2 - 1 = 0$

$x_1 = 1/2$

$x_2 = 1/2$

$x_3 = 0.$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

H:

$$\begin{bmatrix} \frac{\partial}{\partial x_1}\left(\frac{\partial f}{\partial x_1}\right) & \frac{\partial}{\partial x_1}\left(\frac{\partial f}{\partial x_2}\right) & \frac{\partial}{\partial x_1}\left(\frac{\partial f}{\partial x_3}\right) \\ \frac{\partial}{\partial x_2}\left(\frac{\partial f}{\partial x_1}\right) & \frac{\partial}{\partial x_2}\left(\frac{\partial f}{\partial x_2}\right) & \frac{\partial}{\partial x_2}\left(\frac{\partial f}{\partial x_3}\right) \\ \frac{\partial}{\partial x_3}\left(\frac{\partial f}{\partial x_1}\right) & \frac{\partial}{\partial x_3}\left(\frac{\partial f}{\partial x_2}\right) & \frac{\partial}{\partial x_3}\left(\frac{\partial f}{\partial x_3}\right) \end{bmatrix}$$

$$H = \begin{bmatrix} 4 & -2 & 0 \\ -2 & 4 & 0 \\ 0 & 0 & 2 \end{bmatrix} \Rightarrow H-\lambda I = \begin{bmatrix} 4-\lambda & -2 & 0 \\ -2 & 4-\lambda & 0 \\ 0 & 0 & 2-\lambda \end{bmatrix}$$

$$\det(H-\lambda I) = 0 \Rightarrow (4-\lambda)(4-\lambda)(2-\lambda) + 2(-2)(2-\lambda) = 0$$

$$= (2-\lambda)\left[(4-\lambda)^2 - 4\right] = 0$$

$$2-\lambda = 0 \Rightarrow \lambda = 2;$$

$$\lambda^2 - 8\lambda + 16 - 4 = 0 \qquad \lambda^2 - 8\lambda + 12 = 0$$

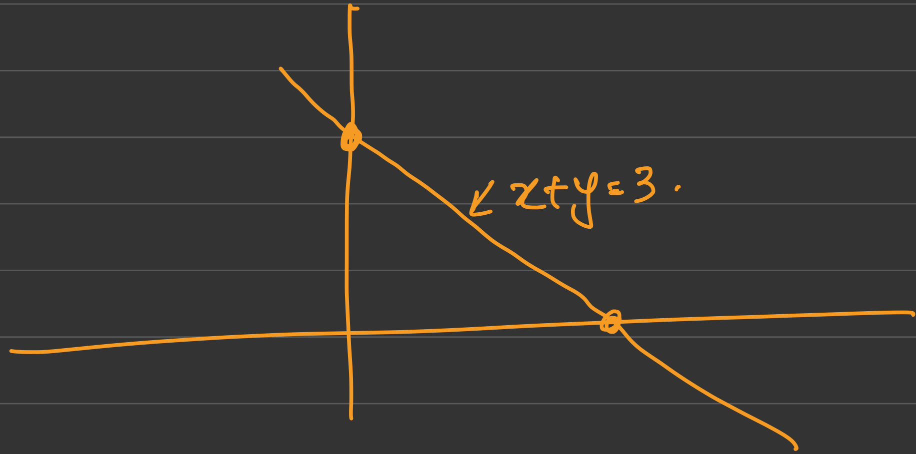
$$(\lambda-6)(\lambda-2) = 0$$

$$\lambda = 2, 6, 2.$$

$\therefore$ H is positive definite.

$\therefore$ f has min at $\begin{pmatrix} 1/2 \\ 1/2 \\ 0 \end{pmatrix}$.

Min $2x^2 + 4y^2 - 3xy$.
Subject to $x + y = 3$.



Lagrangian Multiplier: Converts a constrained optimizn problem to an unconstrained one.

* $f(x,y) = x^2 + 4y^2 - 2x + 8y$ Constraint $x + 2y = 7$.

$x + 2y - 7 = 0$

Lagrangian λ.

$\lambda(x + 2y - 7)$

Objec: $f(x,y) + \lambda g(x,y)$

$= x^2 + 4y^2 - 2x + 8y + \lambda(x + 2y - 7)$

$$\frac{\partial}{\partial x} = 2x - 2 + \lambda$$

$$\frac{\partial}{\partial y} = 8y + 8 + 2\lambda$$

$$2x - 2 + \lambda = 0$$

$$\Rightarrow$$

$$2x + (\lambda - 2) = 0.$$

$$\partial = 8y + (2\lambda + 8) = 0.$$

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

$$y = \frac{-8-2\lambda}{8}$$

Constraint: $x + 2y = 7$

$$\frac{2-\lambda}{2} + \cancel{2}\frac{(-8-2\lambda)}{\cancel{8}\,4} = 7$$

$$\frac{4 - 2\lambda - 8 - 2\lambda}{4} = 7 \Rightarrow -4 - 4\lambda = 28$$

$$-4\lambda = 32$$

$$\lambda = -8$$

$$x = \frac{2+8}{2}$$

$$y = \frac{-8+16}{8}$$

$$(x, y) = (5, 1).$$

Ex: 2 $\underbrace{x_1^2 + x_2^2 - 4x_1 - 4x_2}$ Sub. $x_1 + x_2 = 2$

using Lagrange. Find the min value of the fn

obj: $x_1^2 + x_2^2 - 4x_1 - 4x_2 + \lambda(x_1 + x_2 - 2)$.

$\frac{\partial}{\partial x_1} = 2x_1 - 4 + \lambda$ $= 0$

$\frac{\partial}{\partial x_2} = 2x_2 - 4 + \lambda$ .. $= 0$

$x_1 = \frac{4-\lambda}{2} = x_2$.

$x_1 + x_2 = 2$. $\frac{4-\lambda}{2} + \frac{4-\lambda}{2} = 2$. $(4-\lambda) = 2$

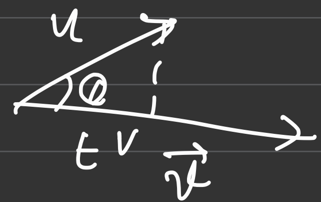
$\lambda = 2$.

$x_1 = x_2 = 1$ $f(1,1) = -6$.

Principal Component Analysis.

→ eigenvectors corresp. to dominant eigenval. of cov matrix

PCA : Projection of higher dimensional data to lower dimensional space.



N- data points → $\vec{y} \in \mathbb{R}^n$.

$t = \dfrac{u \cdot v}{v \cdot v}$.

Obj. Max $x = W^T(y - b)$. Subject to $W^T W = 1$

Prove

9:00 → 10:00

||