1) a) $AA^T = I$

ORTHOGONAL

$A^T = A^{-1}$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} a & c \\ b & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$a^2 + b^2 = 1$

$c^2 + d^2 = 1$

$ac + bd = 0$

$a = 1$

$b = 0$

$c = 0$

$d = 1$

$a = 1/\sqrt{2}$

$b = -1/\sqrt{2}$

$c = 1/\sqrt{2}$

$d = 1/\sqrt{2}$

$$\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

$$\begin{bmatrix} c\theta & -s\theta \\ s\theta & c\theta \end{bmatrix}$$

$$A = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

$$Av = \lambda v$$

$$+3\lambda + \lambda^2 + 2$$

$$|A - \lambda I| = 0$$

$$\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} \frac{1}{\sqrt{2}} - \lambda & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} - \lambda \end{bmatrix}$$

✓ SINGULAR

$$\left(\frac{1}{\sqrt{2}} - \lambda\right)^2 - \left(-\frac{1}{2}\right) = 0$$

$$\frac{1}{2} + \lambda^2 - \frac{2\lambda}{\sqrt{2}} + \frac{1}{2} = 0$$

$\sqrt{-2}$

$$\lambda^2 - \sqrt{2}\lambda + 1 = 0$$

$$\lambda = \frac{\sqrt{2} \pm \sqrt{2-4}}{2} = \frac{1}{\sqrt{2}} \pm \frac{\sqrt{2}i}{2}$$

$$= 1/\sqrt{2} \pm i\, 1/\sqrt{2}$$

$$\lambda_1 = 1/\sqrt{2} + i/\sqrt{2} \qquad \lambda_2 = 1/\sqrt{2} - i/\sqrt{2}$$

$$A - \lambda I = \begin{bmatrix} 1/\sqrt{2} - \lambda & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} - \lambda \end{bmatrix} \qquad A - \lambda_2 I = \begin{bmatrix} i/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & i/\sqrt{2} \end{bmatrix}$$

$$A - \lambda_1 I = \begin{bmatrix} -i/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & -i/\sqrt{2} \end{bmatrix} x = 0$$

$$-i/\sqrt{2}\, x_1 - 1/\sqrt{2}\, x_2 = 0$$

$$1/\sqrt{2}\, x_1 - i/\sqrt{2}\, x_2 = 0$$

$$-ix_1 = x_2$$

$$x_1 = ix_2$$
$$= i(-ix_1)$$
$$= -i^2 x_1$$
$$= x_1$$

$x_2 = 1$

$x_1 = i$

$x_1 = i$

$x_2 = 1$

$$x_1 = \begin{bmatrix} i \\ 1 \end{bmatrix} \mathbin{/\!/}$$

$$\begin{bmatrix} -i \\ 1 \end{bmatrix}$$

$$A - \lambda_2 I = \begin{bmatrix} i/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & i/\sqrt{2} \end{bmatrix}$$

$$ix_1 - x_2 = 0$$

$$x_1 + ix_2 = 0$$

$$x_2 = ix_1$$

$$x_1 = -ix_2$$

$x_1 = 1$

$x_2 = i$

$$x_2 = \begin{bmatrix} 1 \\ i \end{bmatrix} \mathbin{/\!/}$$

→ EIGENVALUES → not real → need not be always

complex conjugates

$$|EV| = 1$$

Evectors came out orthogonal

ii) $AA^T = I$

$A^T = A^{-1}$

$Ax = \lambda x$

$\|Ax\|^2 = \|\lambda x\|^2$

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

$(Ax)^T (Ax)$

$x^T \underbrace{A^T A}\, x$

$x^T x \qquad \|x\|^2$

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

$|\lambda|^2 = 1$

$|\lambda| = 1$ as length > 0 //

iii) $Ax_1 = \lambda_1 x_1$ $\qquad x_1^T A^T A x_1 = |\lambda|^2 \, ||x_1||^2$

$Ax_2 = \lambda_2 x_2$ $\qquad x_2^T A^T A x_2 = |\lambda_2|^2 \, ||x_2||^2$

$$x_1 . x_2 = x_1^{\dagger} x_2 = 0$$

$$\underline{x_1^{\dagger} \cdot x_2} = x_1^{\dagger} . I . x_2$$

$$= x_1^{\dagger} . \underline{A^T A} . x_2$$

$$= (Ax_1)^{\dagger} . Ax_2$$

$$= (\lambda_1 x_1)^{\dagger} . \lambda_2 x_2$$

$$= \bar{\lambda}_1 \lambda_2 (x_1^{\dagger} x_2)$$

$$x_1^{\dagger} x_2 (\bar{\lambda}_1 \lambda_2 - 1) = 0$$

$\underline{\lambda_1 = \lambda_2}$

$\underline{x_1^T x_2 = 0}$ or $\underline{\bar{\lambda}_1 \lambda_2 = 1}$

$(a_1 - ib)(c + id) = 1$

We know $|\lambda| = 1$

$|\bar{\lambda}_1 \lambda_2| =$

$ac + bd + i(ad - bc) = 1$

$ad - bc = 0$

$ac + bd = 1$

$$\lambda_1 = \cos\theta + i\sin\theta$$

$$\lambda_2 = \cos\phi + i\sin\phi$$

$$\bar{\lambda}_1 \lambda_2$$

$$(\cos\theta - i\sin\theta)(\cos\phi + i\sin\phi) = 1$$

$$\cos\theta\cos\phi - i\sin\theta\cos\phi + i\cos\theta\sin\phi + \sin\theta\sin\phi = 1$$

$$\sin\phi\cos\theta - \sin\theta\cos\phi = 0$$

$$\sin\theta\sin\phi + \cos\theta\cos\phi = 1$$

$$\sin(\phi - \theta) = 0$$

$$\cos(\phi - \theta) = 1$$

$$\phi - \theta = n\pi$$

$$0 \leq \phi - \theta \leq 2\pi$$

$$\phi = \theta + n\pi \;]$$

so $\nearrow\nearrow$ or $\nearrow\swarrow$

$\lambda_1 = \lambda_2$ or $\lambda_1, \lambda_2 \to$ complex conjugates

If real $\Rightarrow \lambda_1 = \lambda_2$ //

$|\lambda_1|^2 = 1 \quad |\lambda_2|^2 = 1$

$a^2 + b^2 = 1$

$c^2 + d^2 = 1$

$a^2c^2 + b^2d^2 + 2abcd = 1$

$\underline{a^2c^2} + a^2d^2 + b^2c^2 + \underline{b^2d^2} = 1$

$a^2d^2 + b^2c^2 + 2abcd = 0$

$(ad + bc)^2 = 0$

$ad + bc = 0$

$ad = 0 \quad bc = 0$

$a + ib$

$c + id$

$\bar{\lambda}_1 \lambda_2 = 1$

$\bar{\lambda}_2 \lambda_1 = 1$

$|\lambda_1| = |\lambda_2| = 1$

To Prove

$\lambda_1 = \lambda_2$

iv) It preserves inner product and keeps distance.

eg: Reflection / Rotation

1b) i) $A = U \Sigma V^T$

$$AA^T = U \Sigma V^T (U \Sigma V^T)^T$$

$$= U \Sigma \underbrace{V^T V}_{\text{orthogonal}} \Sigma^T U^T$$

$$= U (\Sigma \Sigma^T) U^T$$

$$\boxed{X = S \Lambda S^{-1}}$$

compare

$S = U$

$S^{-1} = U^T$

↳ makes sense AA^T is always symmetric positive definite

So U are the eigenvectors of AA^T

$$\Sigma \Sigma^T = \begin{bmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_n^2 \end{bmatrix}$$

The eigenvalues of AA^T = Square of the singular values of A.

$$A^T A = (U \Sigma V^T)^T U \Sigma V^T$$

$$= V \Sigma^T \underline{U^T U} \Sigma V^T$$

V → eigenvectors of $A^T A$

The eigenvalues of $A^T A$ = Square of the singular values of A.

c) i) at most n distinct

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

FALSE

2D space

eigenvalues = 1, 1

ii) $Ax_1 = \lambda_1 x_1$

FALSE

$Ax_2 = \lambda_2 x_2$

$A(x_1 + x_2) = Ax_1 + Ax_2$

$= \lambda_1 x_1 + \lambda_2 x_2$

if $\lambda_1 = \lambda_2$

$x_1 + x_2$ is eigenvector

$\lambda_1 \neq \lambda_2$

if $A(x_1 + x_2) = \mu(x_1 + x_2)$

$\lambda_1 x_1 + \lambda_2 x_2 = \mu x_1 + \mu x_2$

$x_1(\lambda_1 - \mu) + x_2(\lambda_2 - \mu) = 0$

we know if $\lambda_1 \neq \lambda_2$

then x_1, x_2 are linearly independent

$ax_1 + bx_2 = 0$ prove $a = b = 0$

$\times A$

$Aax_1 + Abx_2 = 0$

$a\lambda_1 x_1 + b\lambda_2 x_2 = 0$ - 1

$\times \lambda_1$

$$a\lambda_1 x_1 + b\lambda_1 x_2 = 0 \quad - ②$$

② − ①

$$b x_2 (\lambda_1 - \lambda_2) = 0$$

not 0   not 0

$b = 0$

III^ly $a = 0$

So $(\lambda_1 - \mu) = \lambda_2 - \mu = 0$

$\lambda_1 = \lambda_2$ CONTRADICTION

iii) $x^T A x \geq 0$ for all x TRUE

$A x_i = \lambda_i x_i$ $x_i \to$ EV

$$x^T \lambda_i x \geq 0$$

$$\lambda_i x_i^T x \geq 0$$

$$\lambda_i \|x_i\|^2 \geq 0$$

$\hookrightarrow \geq 0$

so $\lambda_i \geq 0$

iv) TRUE

'distinct'

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Rank = 2

v) $Ax_1 = \lambda x_1$

$Ax_2 = \lambda x_2$

$$A(x_1 + x_2) = \lambda(x_1 + x_2)$$

TRUE

PROBABILITY:

2) a) i) H60 H50

$$P(H50 \mid T) = \frac{P(T \mid H50) \times P(H50)}{P(T)}$$

$$= \frac{0.5 \times 1/2}{0.5 \times 1/2 + 0.4 \times 1/2}$$

$$= \frac{0.5}{0.9} = 5/9$$

ii) THHH

$$P(H50 \mid THHH) = \frac{P(THHH \mid H50) \cdot \cancel{P(H50)}}{P(THHH \mid H50) \cdot \cancel{P(H50)} + P(THHH \mid H60) \cdot \cancel{P(H60)}}$$

$$= \frac{(1/2)^4}{(1/2)^4 + \left(\frac{2}{5}\right)\left(\frac{3}{5}\right)^3}$$

iii) $$P(H50 \mid 9H1T) = \frac{P(9H1T \mid H50) \cdot \cancel{P(H50)}}{P(9H1T \mid H50) \cdot \cancel{P(H50)} + P(9H1T \mid H55) \cdot \cancel{P(H55)} + P(9H1T \mid H60) \cdot \cancel{P(H60)}}$$

$$= \frac{(0.5)^9 (0.5)^1}{(0.5)^{10} + (0.55)^9 (0.45) + (0.6)^9 (0.4)}$$

b) $P(+ \mid P) = 99/100$ $\quad P(+ / N) = 10/100 = 1/10$

$P(- \mid P) = 1/100$ $\quad P(- / N) = 9/10$

$P(N) = 99/100$ $\quad P(P) = 1/100$

$$P(P \mid +) = \frac{P(+ \mid P) \cdot P(P)}{P(+ \mid P) \cdot P(P) + P(+ \mid N) \cdot P(N)}$$

$$= \frac{99/100 \times 1/100}{99/100 \times 1/100 + 10/100 \times 99/100}$$

$$= \frac{99}{99 + 990}$$

$$= \frac{99}{1089}$$

The population has majority -ve.

c) $AE(x) + b$

d) $E[(x - E(x))(x - E(x))^T]$

$E[(Ax + \cancel{b} - AE(x) - \cancel{b})(Ax + \cancel{b} - AE(x) - \cancel{b})^T]$

$E[\underbrace{A(x - E(x)) \quad (x - E(x))^T A^T}]$

$A \operatorname{Cov}(x) A^T$

MULTIVARIATE DERIVATES:

a) $x \in \mathbb{R}^n$

$y \in \mathbb{R}^m$

$A \in \mathbb{R}^{n \times m}$

$$\nabla_x \, x^T A y$$

If $y = Ax$

$$\frac{\partial y}{\partial x} = A$$

$$\Rightarrow \quad y_i = \sum_{k=1}^{n} a_{ik} x_k$$

$$\frac{\partial y_i}{\partial x_j} = a_{ij}$$

$$\frac{\partial y}{\partial x} = A$$

$$\nabla_x \; \underset{1\times n}{x^T} \; \underset{n\times m}{A} \; \underset{m\times 1}{y}$$

$\alpha = x^T A y \rightarrow$ scalar

$$\alpha = \alpha^T = (x^T A y)^T = y^T A^T x$$

$$\nabla_x x^T A y = \nabla_x \underbrace{y^T A^T}_{\text{of form } Ax} x$$

$$= \underset{1\times m}{y^T} \; \underset{m\times n}{A^T}$$

OR

$$z = x^T A y$$

$$= \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \underset{m\times n}{\begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix}} \underset{n\times 1}{\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}}$$

$$= \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} a_{11}y_1 + a_{12}y_2 + y_{1m}y_m \\ \vdots \qquad\qquad \vdots \\ a_{n1}y_1 + a_{n2}y_2 + y_{nm}y_n \end{bmatrix}$$

$$= (a_{11}y_1 + a_{12}y_2 + \cdots + a_{1m}y_m)x_1 +$$

$$(a_{21}y_1 + a_{22}y_2 + \cdots + a_{2m}y_m)x_2 +$$

$$\vdots$$

$$(a_{n1}y_1 + a_{n2}y_2 + \cdots + a_{nm}y_n)x_n$$

$$z = \sum_{i=1}^{n} (a_{i1}y_1 + a_{i2}y_2 + \cdots + a_{im}y_m)x_i$$

$$z = \sum_{i=1}^{n} \sum_{j=1}^{m} (a_{ij}y_j)x_i$$

$$\frac{\partial z}{\partial x_i} = \sum_{j=1}^{m} a_{ij}y_j \qquad \underset{n\times m}{A}\underset{m\times 1}{y}$$

$$\nabla_x x^T A y = Ay$$

b) $\nabla_y \; x^T A y$ $\qquad\qquad\qquad A^T x$

$$\frac{\partial z}{\partial y_j} = \sum_{i=1}^{n} (a_{ij} x_i)$$

$$\downarrow \qquad \searrow$$

$n \times m \qquad n \times 1$

$$\frac{\partial z}{\partial y} = \begin{bmatrix} a_{11}x_1 + a_{21}x_2 + \cdots + a_{n1}x_n \\ a_{12}x_1 + a_{22}x_2 + \cdots + a_{n2}x_n \\ \vdots \\ a_{1m}x_1 + a_{2m}x_2 + \cdots + a_{nm}x_n \end{bmatrix}$$

$$\begin{bmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & & & \\ \vdots & & & \\ a_{1m} & \cdots & & a_{nm} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$= A^T x$$

c) $z = \sum_{i=1}^{n} \sum_{j=1}^{m} (a_{ij} y_j) x_i$

$$\frac{\partial z}{\partial a_{ij}} = x_i y_j$$

$$\frac{\partial z}{\partial A} = \underset{n \times 1}{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}} \underset{1 \times m}{\begin{bmatrix} y_1 & \cdots & y_m \end{bmatrix}}$$

$$= xy$$

d) $f = \underbrace{x^T A x}_{g} + \underbrace{b^T x}_{h}$

$$\begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \vdots \\ a_{n1} & \cdots & & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$g_{ij} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$$

If $i = j$

$$\frac{\partial g}{\partial x_i} = 2a_{ij} x_i$$

$i \neq j$

$$\frac{\partial g}{\partial x_i} = \sum_{i=1}^{n} a_{ij} x_j$$

$$\frac{\partial g}{\partial x_i} = 2a_{ij} x_i \quad \sum_{\substack{i=1 \\ i \neq j}}^{n} a_{ij} x_j + \sum_{\substack{j=1 \\ i \neq j}}^{n} a_{ji} x_j$$

$$= \sum_{j=1}^{n} (a_{ij} + a_{ji}) x_j$$

$$= \sum_{j=1}^{n} a_{ij} x_j + \sum_{j=1}^{n} a_{ji} x_j$$

$$\frac{\partial g}{\partial x} = \begin{bmatrix} a_{11} & a_{12} \cdots & a_{1n} \\ \vdots & & \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$+ \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ \vdots & & & \\ a_{1n} & \cdots & & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

$$= Ax + A^T x$$

$$h = b^T x$$

$b^T \to 1 \times n$, $x \to n \times 1$

$$b \to n \times 1$$

$$\begin{bmatrix} b_1 & b_2 & \cdots & b_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$h = b_1 x_1 + \cdots + b_n x_n$$

$$= \sum_{i=1}^{n} b_i x_i$$

$$\frac{\partial h}{\partial x_i} = b_i$$

$$\frac{\partial h}{\partial x} = b$$

$$\frac{\partial f}{\partial x} = \nabla_x f = Ax + A^T x + b$$

e) $f = tr(AB)$

$\nabla_A f$

$A \to n \times m$

$B \to m \times n$

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ \vdots & & & \\ a_{n1} & \cdots & & a_{nm} \end{bmatrix} \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + \cdots + a_{1m}b_{m1} & & \\ & a_{21}b_{12} + a_{22}b_{21} + \cdots + a_{2m}b_{m2} & \\ & \ddots & \\ & & a_{n1}b_{1n} + a_{n2}b_{2n} + \cdots + a_{nm}b_{mn} \end{bmatrix}$$

$$\sum_{i=1}^{m} a_{1i} b_{i1}$$

$$\sum_{j=1}^{n} \sum_{i=1}^{m} a_{ji} b_{ij} \qquad \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij} b_{ji}$$

$$\frac{\partial f}{\partial a_{ij}} = b_{ji}$$

$$\frac{\partial f}{\partial A} = B^T \quad //$$

4) $\overset{q\times 1}{\hat{y}} = \overset{q\times p}{W} \overset{p\times 1}{x}$

$\underset{p\times 1}{x} \longrightarrow \underset{q\times 1}{y}$ $\qquad W \to q\times p$

$$L = \frac{1}{2}\sum_{i=1}^{n} \|y^{(i)} - Wx^{(i)}\|^2$$

$1\times m$

$$= \frac{1}{2}\sum_{i=1}^{n} (y_i - Wx_i)^T (y_i - Wx_i)$$

$$\begin{bmatrix} y_{11} \\ y_{12} \\ \vdots \\ y_{1q} \end{bmatrix} - \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1p} \\ w_{21} & & & \\ \vdots & & & \\ w_{q1} & \cdots & & w_{1qp} \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{12} \\ \vdots \\ x_{1p} \end{bmatrix}$$

$$L = \frac{1}{2}\sum_{i=1}^{n} (y_i - Wx_i)^T (y_i - Wx_i)$$

$$(y_i^T - x_i^T W^T)(y - Wx_i)$$

$y_i - Wx_i$ when applied to all i

$$Y - WX = \begin{bmatrix} y_{11} & y_{21} \cdots & y_{n1} \\ y_{12} & y_{22} & \vdots \\ \vdots & & \\ y_{1q} & & y_{nq} \end{bmatrix}_{q\times n} - \begin{bmatrix} & & \\ & W & \\ & & \end{bmatrix}_{q\times p} \begin{bmatrix} x_{11} & x_{21} & \cdots & x_{n1} \\ x_{12} & & & \vdots \\ \vdots & & & \\ x_{1p} & & & x_{np} \end{bmatrix}_{p\times n}$$

$$Y - WX$$

$$\downarrow$$

$$(Y - WX)^T (Y - WX)$$

$$\begin{bmatrix} (y_1 - Wx_1)^T (y_1 - Wx_1) & (y_1 - Wx_1)^T (y_2 - Wx_2) & \cdots \\ \vdots & (y_2 - Wx_2)^T (y_2 - Wx_2) & \\ \vdots & & \ddots \\ & & (y_n - Wx_n)^T (y_n - Wx_n) \end{bmatrix}$$

we need the sum of diagonals

$$L = \frac{1}{2} \operatorname{trace} (Y - WX)^T (Y - WX)$$

$$= \frac{1}{2} \operatorname{tr} (Y^T - X^T W^T)(Y - WX)$$

$$L = \frac{1}{2} \, tr \, (y^T y - x^T W^T y - y^T W x + x^T W^T W x)$$

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

0

$$\frac{\partial L}{\partial W} = \frac{\partial}{\partial W} tr(y^T y) + \frac{\partial}{\partial W} Tr(x^T W^T W x) - \frac{\partial}{\partial W} tr(x^T W^T y) - \frac{\partial}{\partial W} tr(y^T W x)$$

T

$tr(A^T) = tr(A)$

$tr(ABC) =$

$tr(CAB)$

$$+\frac{\partial}{\partial W} Tr(x x^T W^T W)$$

$$-\frac{2\partial}{\partial W} Tr(y^T W x)$$

$$= -\frac{2\partial}{\partial W} Tr(W x y^T)$$

$$= -2(y x^T)$$

$$\frac{+\partial}{\partial W}\left(W x x^T W^T\right)$$

$A = x x^T$

$(x x^T)^T$

$W A^T + W A$

$(x^T)^T x^T$

$W(x x^T) + W(x x^T)$

$+ 2W x x^T$

$$2W x x^T = 2 y x^T$$

$$W = y x^T (x x^T)^{-1} \;//$$