Fischer's LDA:

↳ Dimensionality reduction
↳ Labelled data (unlike PCA)

Need: Direction to amplify class difference

↳ PCA won't work as it will increase overlap

↳ Using maximised (squared) diff

$$W = \arg\max (m_1 - m_2)^2$$

as there is still overlap (but lesser overlap)
large scatter from each class
But not too bad

So minimise projected scatter within each class $s_1^2 + s_2^2$
and max. (squared) distance b/w projected means $(m_1 - m_2)^2$

w from LDA

Maximise fischer index:

$$J(w) = \frac{(m_1 - m_2)^2}{s_1^2 + s_2^2}$$

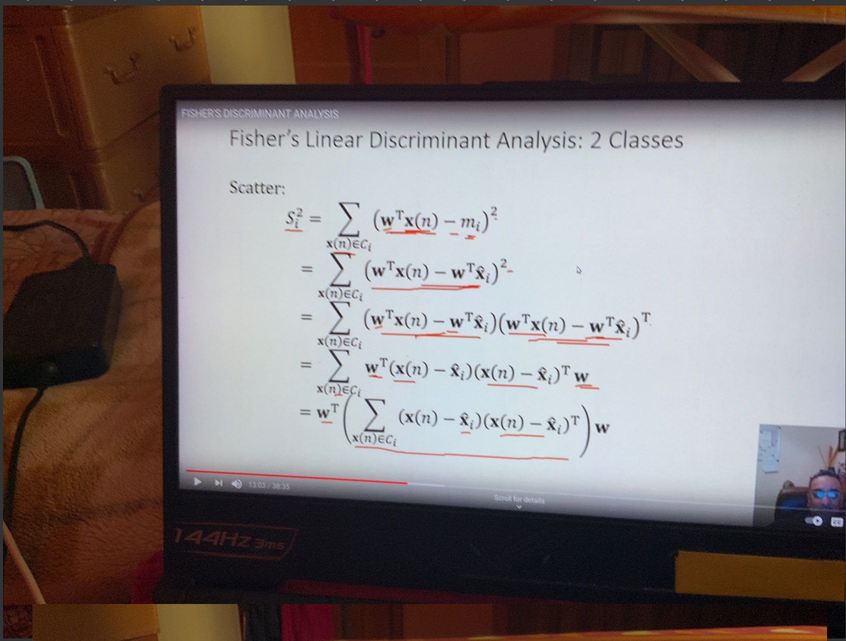
Direction of projection: w

Two classes C_i, $i = 1, 2$
No. of samples: $N_i = (C_i)$
Projection of samples: $w^T x(n)$

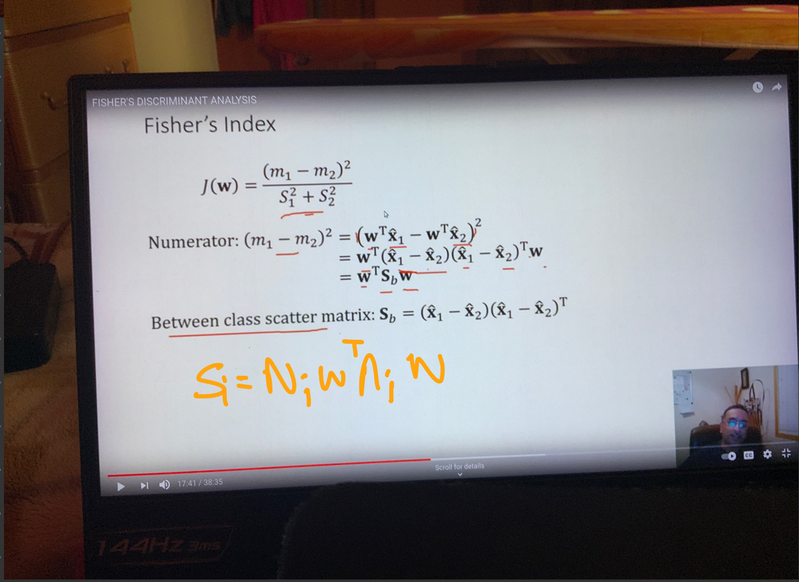
Means along proj.: $m_i = w^T \hat{x}_i$

Variance:

$$A_i = \frac{1}{N_i} \sum_{x(n) \in C_i} (x(n) - \hat{x}_i)(x(n) - \hat{x}_i)^T$$



Variation of projected points along direction is $w w_j^T x(n)$



Denominator:

$N_1 w^T \Lambda_1 w + N_2 w^T \Lambda_2 w$

$w^T (N_1 \Lambda_1 + N_2 \Lambda_2) w$

$w^T S_w w$

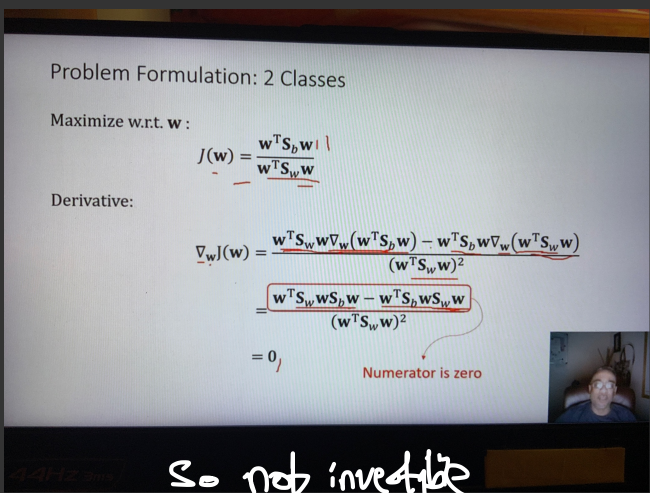
Within class matrix

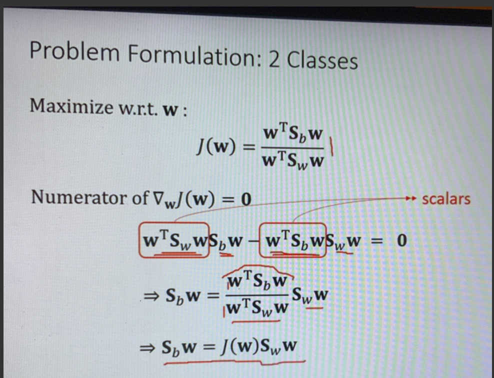
Maximise:

$$J(w) = \frac{w^T S_b w}{w^T S_i w}$$

$\rightarrow (\hat{x}_1 - \hat{x}_2)(\hat{x}_1 - \hat{x}_2)^T$

↳ $N_i w^T \Lambda_i w$





So not invertible

Each sample with high dimension

Solution: $S_b w = J(w) S_w w$

$\Rightarrow S_b w = \lambda S_w w$

Generalised eigenvalue-vector $\lambda_i v_i$:

$w = v_1, J(w) = \lambda_1$ as $Av = \lambda B v$

↳ Not taking inverse as $\text{rank}(S_w) \le N$ [No. of sample as in summation]

FISHER'S DISCRIMINANT ANALYSIS

Reminder

$D \times D$ Covariance matrix: $\mathbf{\Lambda}$

$$\mathbf{\Lambda} = \frac{1}{N}\sum_{n}(\mathbf{x}(n) - \hat{\mathbf{x}})(\mathbf{x}(n) - \hat{\mathbf{x}})^{\mathrm{T}}$$

Rank of each term: $(\mathbf{x}(n) - \hat{\mathbf{x}})(\mathbf{x}(n) - \hat{\mathbf{x}})^{\mathrm{T}}$ is 1

Rank of summation: $\mathbf{\Lambda} \leq N$

Within class scatter matrix: $\mathbf{S}_w = N_1\mathbf{\Lambda}_1 + N_2\mathbf{\Lambda}_2$

Rank of $\mathbf{S}_w \leq N$

$D \times D$

S DISCRIMINANT ANALYSIS

Fisher's Linear Discriminant Analysis: C Classes

Between class scatter matrix:

$$\mathbf{S}_b = \sum_{i=1}^{C} N_i(\hat{\mathbf{x}}_i - \hat{\mathbf{x}})(\hat{\mathbf{x}}_i - \hat{\mathbf{x}})^{\mathrm{T}}$$

$\boxed{\operatorname{rank}(\mathbf{S}_b) \leq C - 1}$

Within class scatter matrix:

$$\mathbf{S}_w = \sum_{i=1}^{C}\sum_{\mathbf{x}(n)\in C_i}(\mathbf{x}(n) - \hat{\mathbf{x}}_i)(\mathbf{x}(n) - \hat{\mathbf{x}}_i)^{\mathrm{T}}$$

Generalized eigenvalue-eigenvector: $\mathbf{S}_b\mathbf{w} = \lambda\mathbf{S}_w\mathbf{w}$

Cannot seek more than $\boxed{C - 1}$ eigenvalues!

Rest will be 0

FISHER'S DISCRIMINANT ANALYSIS

Fisher's Linear Discriminant Analysis: C Classes

Projection matrix: $\mathbf{W} = [\mathbf{w}_1\ \mathbf{w}_2 \ldots \mathbf{w}_{C-1}]$

Maximize w.r.t. $\mathbf{w}$: $J(\mathbf{W}) = \frac{|\mathbf{W}^{\mathrm{T}}\mathbf{S}_b\mathbf{W}|}{|\mathbf{W}^{\mathrm{T}}\mathbf{S}_w\mathbf{W}|}$ (Generalized Fisher index)

Generalized eigenvalue-eigenvector: $\mathbf{S}_b\mathbf{v} = \lambda\mathbf{S}_w\mathbf{v}$

Solution: $\mathbf{W} = [\mathbf{v}_1\ \mathbf{v}_2 \ldots \mathbf{v}_{C-1}]$

$\boxed{\text{Cannot seek more than } C - 1 \text{ eigenvalues!}}$

Fundamental limitation of

Face Recognition

Faces: $\mathbf{F}(n), n = 1,2,\ldots N$

$l \times b$ matrix of pixels $= l \cdot b \times 1 = D \times 1$ column vectors: $\mathbf{x}(n)$

No. of classes: C

Approach:

1. Reduce from D dims to $N - C$ using PCA
2. Reduce from $N - C$ dims to $C - 1$ dims using LDA
3. Use nearest neighbor approach to classify in $C - 1$ dims

FISHER'S DISCRIMINANT ANALYSIS

Face Recognition

1. Reduce from D dims to $N - C$ using PCA

Compute $D \times D$ total scatter matrix $\mathbf{S}_t$

$$\mathbf{S}_t = \sum_{n}(\mathbf{x}(n) - \hat{\mathbf{x}})(\mathbf{x}(n) - \hat{\mathbf{x}})^{\mathrm{T}} = N\mathbf{\Lambda}$$

Construct $\mathbf{W}_{\mathrm{PCA}}$ using $N - C$ eigenvectors of $\mathbf{S}_t$

$$\mathbf{y}(n) = \mathbf{W}_{\mathrm{PCA}}^{\mathrm{T}}\mathbf{x}(n)$$

FISHER'S DISCRIMINANT ANALYSIS

Face Recognition

2. Reduce from $N - C$ dims to $C - 1$ using LDA

Compute $(N - C) \times (N - C)$ matrices $\mathbf{S}_b, \mathbf{S}_w$

$$\mathbf{S}_b = \sum_{i=1}^{C} N_i(\hat{\mathbf{x}}_i - \hat{\mathbf{x}})(\hat{\mathbf{x}}_i - \hat{\mathbf{x}})^{\mathrm{T}}$$

$$\mathbf{S}_w = \sum_{i=1}^{C}\sum_{\mathbf{x}(n)\in C_i}(\mathbf{x}(n) - \hat{\mathbf{x}}_i)(\mathbf{x}(n) - \hat{\mathbf{x}}_i)^{\mathrm{T}}$$

Construct $\mathbf{W}_{\mathrm{LDA}}$ using $C - 1$ generalized eigenvectors of $\mathbf{S}_b, \mathbf{S}_w$

$$\mathbf{z}(n) = \mathbf{W}_{\mathrm{LDA}}^{\mathrm{T}}\mathbf{x}(n)$$

$(N-C)(C-1)$

FISHER'S DISCRIMINANT ANALYSIS