- Linear transformation: translate data from R^d to $R^{d'}$

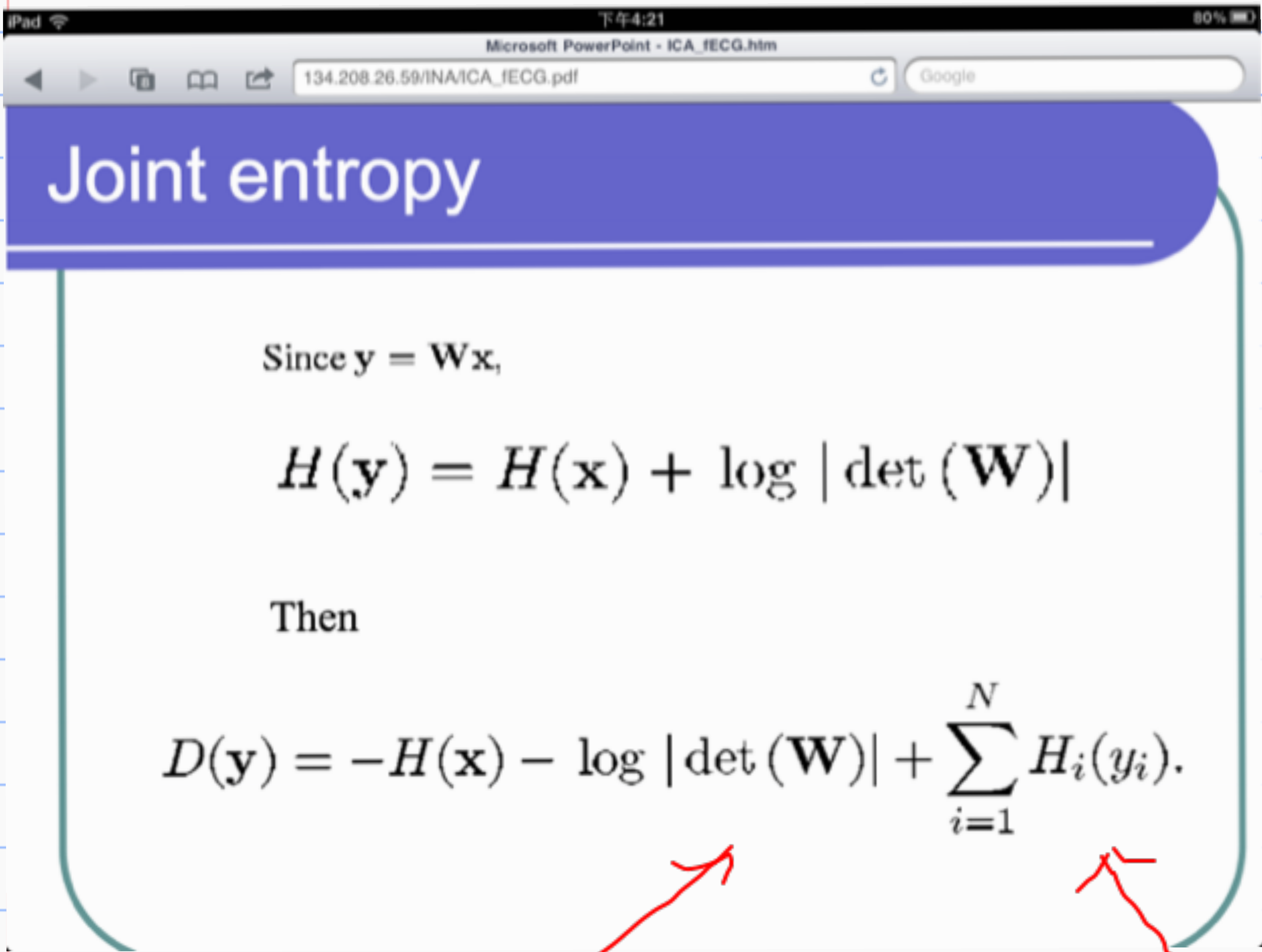
- d == d' : linear transformation for data rotation

- Let B a matrix for linear transformation and b_i^T denote the ith row of B
- $z = Bx$ Matrix
- $z_I = b_i^T x$

- z_i is a projected component
- Orthogonal relation versus independent relation among projected components

- ICA insists on extracting projected components that are statistically independent.
- If B is invertible, the criteria of ICA are respectively minimal volume of a rotated cube spanned by rows of matrix B and maximal individual entropies of extracted components.

ICA



Maximal Volume of a rotated cube

Maximal marginal entropies

- Rows of matrix B are not necessary orthogonal to one another

- Principle component analysis
- b_i^T and b_j^T are orthogonal
- Their inner product is zero
- maximal variance of z_i

- algorithm
- Let $\{x_t\}$ have zero mean
- maximal norm of $b^T x$ provides the criterion of determining the 1st pc
- $\| b \| = 1$

$$\mathbf{w}_1 = \underset{\|\mathbf{w}\|=1}{\arg\max}\, \operatorname{Var}\{\mathbf{w}^T\mathbf{X}\} = \underset{\|\mathbf{w}\|=1}{\arg\max}\, E\left\{\left(\mathbf{w}^T\mathbf{X}\right)^2\right\}$$

(See arg max for the notation.) With the first $k - 1$ components, the kth compo
principal components from **X**:

$$\hat{\mathbf{X}}_{k-1} = \mathbf{X} - \sum_{i=1}^{k-1} \mathbf{w}_i \mathbf{w}_i^T \mathbf{X}$$

and by substituting this as the new data set to find a principal component in

$$\mathbf{w}_k = \underset{\|\mathbf{w}\|=1}{\arg\max}\, E\left\{\left(\mathbf{w}^T\hat{\mathbf{X}}_{k-1}\right)^2\right\}.$$

PCA is equivalent to empirical orthogonal functions (EOF), a name which is us

An autoencoder neural network with a linear hidden layer is similar to PCA. Up
neurons in the hidden layer will form a basis for the space spanned by the first
technique will not necessarily produce orthogonal vectors.

$$E_1(b) = \sum_t \|b^T x_t\|^2$$

$$= b^T \left(\sum_t x_t x_t^T \right) b$$

$$= b^T X b$$

projection norm

- b is an Eigen vector of X
- X sums up $x_t x_t^T$ over t

$$E_2 = b^T b = \|b\|^2$$

$$E = E_1 - \lambda E_2$$

- maximal projection norm
- minimal length of b

$$E = b^T X b - \lambda b^T b$$

$$\frac{dE}{db} = Xb - \lambda b = 0$$

$$\Rightarrow Xb = \lambda b$$

b is an Eigen vector
of X

- RBF

- SOM:

Simplified perceptron
Rosenblatt's general perceptron: