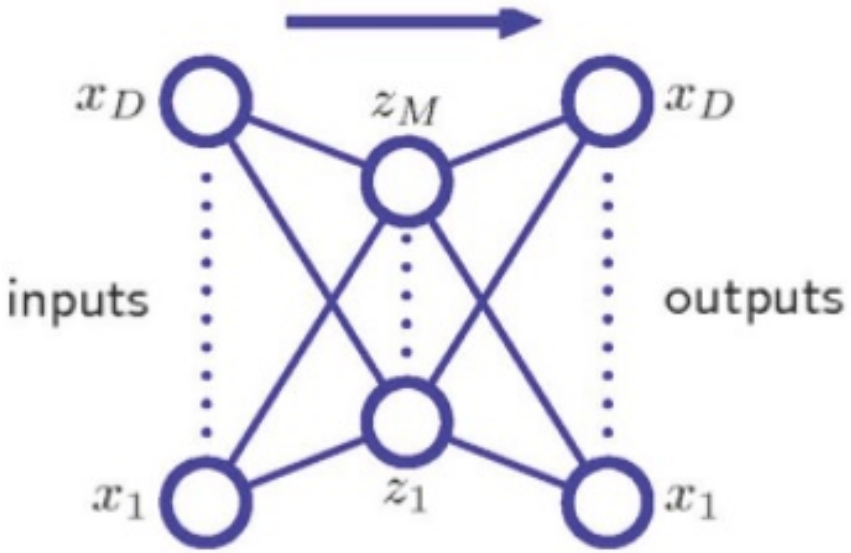
Autoencoder (PCA)

COMP4211

THE DEPARTMENT OF
COMPUTER SCIENCE & ENGINEERING
計算機科學及工程學系

Autoencoder

- neural network whose outputs are its own inputs



- $\mathbf{z} = f(\mathbf{W}\mathbf{x}), \hat{\mathbf{x}} = g(\mathbf{V}\mathbf{z})$
- goal: minimize reconstruction error
  - $\min_{\mathbf{W},\mathbf{V}} \sum_{n=1}^{N} \|\mathbf{x}^{(n)} - \hat{\mathbf{x}}^{(n)}\|^2$
- if f and g are linear
  - $\min_{\mathbf{W},\mathbf{V}} \sum_{n=1}^{N} \|\mathbf{x}^{(n)} - \mathbf{V}\mathbf{W}\mathbf{x}^{(n)}\|^2$
- optimal solution is PCA