Advanced Machine Learning, 8 Oct 2019

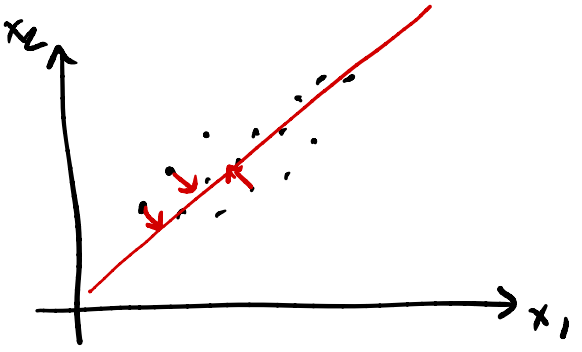
Autoencoder

Nonlinear generalization of PCA

PCA - Principal Component Analysis

Why? Dimension reduction

Example 2D x_1, x_2



Want to reduce to a single dimension

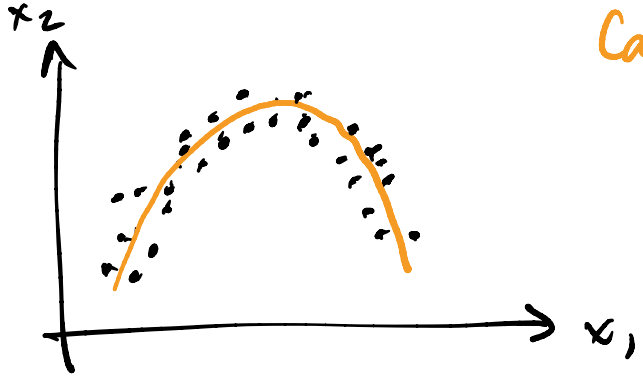
$$h = \beta_1 x_1 + \beta_2 x_2$$

Orthogonal projection onto the line

Limitation - PCA uses linear relationship

- Project onto a smaller dimension hyperplace

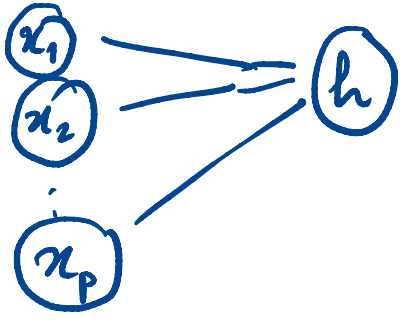
What if, instead, points are as follows?



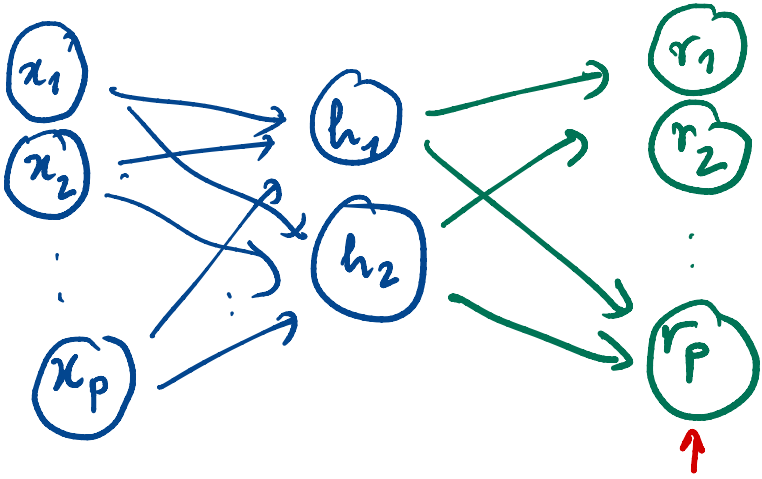
Can we find a curve

Neural network

$$h = f(x_1, x_2) = \sigma(b + x^T w)$$



Can have more than one h

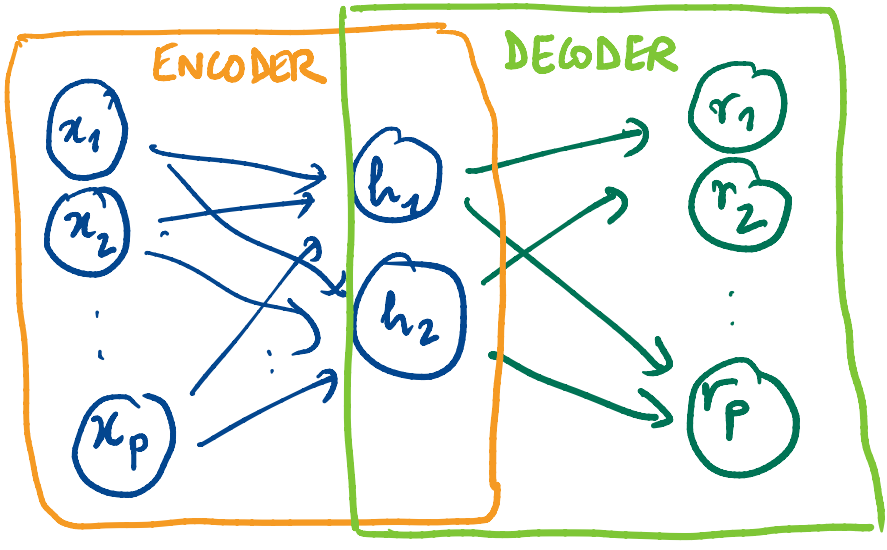


Can interpret as

"Zip" $x_1, \ldots, x_p$ using $h_1, h_2 \ldots$ $x_1, \ldots, x_p$

Recover $x_1, \ldots, x_p$ from h_1, h_2

$r_i \neq x_i$, else nothing achieved

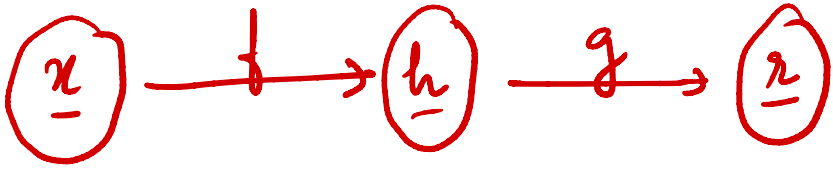


No target variable - unsupervised learning

Undercomplete: Output dimension lower than input

Overcomplete: Output > Input dim

How do we train the weights?



$h = f(x)$, $r = g(h)$

Define loss function

$$L(x, g(f(x)))$$

Detour

$x_1 \dots x_n$ from $N(\mu, \sigma)$

How to estimate μ, σ^2?

Maximize likelihood

$$L(\mu, \sigma \mid \underline{x}) = \prod_{x_i} f(x_i) = \prod \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x_i - \mu}{\sigma}\right)^2}$$

$$-\log L = +\frac{1}{2\sigma^2}\left[\Sigma (x_i - \mu)^2\right] + c$$

square error

Negative log likelihood = Loss function

Focus on estimating h well — want dimension reduction

$$L(\underline{x}, g(h)) + \Omega(h)$$ ← Penalty for h

Suppose we choose square error loss for

$L(\underline{x}, g(h))$

$\Omega(h)$ - prefer sparse encoder - several weights map to 0

2 choices - Laplace (aka lasso)

Student-t

Regularization extremely important for auto encoder

Tuning parameter λ

$$p(h_i) = \frac{\lambda}{2} e^{-\lambda |h_i|}$$

$$\Omega(h) = \lambda \sum_i |h_i|$$

PCA is a special case

- single node, activation fn = identity (not sigmoid, RELU etc)

Deep autoencoder

Multiple layers - benefits similar to deep neural networks

Learning Manifolds with autoencoders - Goodfellow book, 14.6 (p 515)

Non linear surface → manifold

Major success of auto encoder - information retrieval

Key word search - autoencoder hashes the keywords