

COMS30035, Machine learning: Probabilistic PCA

James Cussens

`james.cussens@bristol.ac.uk`

Department of Computer Science, SCEEM
University of Bristol

October 28, 2020

Agenda

- ▶ Probabilistic PCA

Probabilistic PCA

- ▶ Basic idea: reformulate PCA as the maximum likelihood solution to a latent variable model.
- ▶ Unlike with mixtures of Gaussians the latent variable here is continuous.
- ▶ That's why PCA is in the *Continuous Latent Variables* chapter of Bishop.

The PPCA model

The latent variable \mathbf{z} has a zero-mean unit-covariance Gaussian distribution:

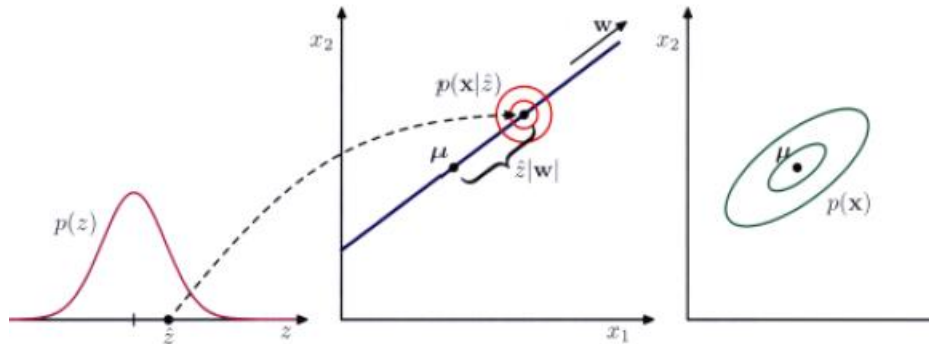
$$p(\mathbf{z}) = \mathcal{N}(\mathbf{z}|\mathbf{0}, \mathbf{I}) \quad (1)$$

The distribution of the observed data conditional on this latent variable is another Gaussian:

$$p(\mathbf{x}|\mathbf{z}) = \mathcal{N}(\mathbf{x}|\mathbf{W}\mathbf{z} + \boldsymbol{\mu}, \sigma^2\mathbf{I}) \quad (2)$$

- ▶ So the parameters (to learn) are: \mathbf{W} , $\boldsymbol{\mu}$ and σ^2 .
- ▶ \mathbf{W} is an $D \times M$ matrix where D is the dimension of the data and M is the dimension of the PCA space.

The generative view of PPCA (Bishop Fig 12.9)



Maximum likelihood for PPCA

- ▶ The good (if unsurprising) news is that the MLE parameters for PPCA can be computed exactly in closed form [Bis06, §12.2.1].
- ▶ And can be ‘read off’ from the (M first) eigenvectors and eigenvalues of the sample covariance matrix.
- ▶ The MLE estimate for μ is just the sample mean.
- ▶ We might still resort to EM if the sample covariance matrix is huge, or if we have to deal with missing values in the data.

Why PPCA?

- ▶ Choosing M : since we now have a likelihood, we can use cross-validation or a Bayesian approach with a special prior on \mathbf{W} .
- ▶ We can make connections to closely related models like *factor analysis* (which is just a small generalisation of PPCA.)
- ▶ We can generate data from a given PPCA model.

Now do the quiz!

Yes, please do the quiz for this lecture on Blackboard!



Christopher M. Bishop.

Pattern Recognition and Machine Learning.

Springer, 2006.