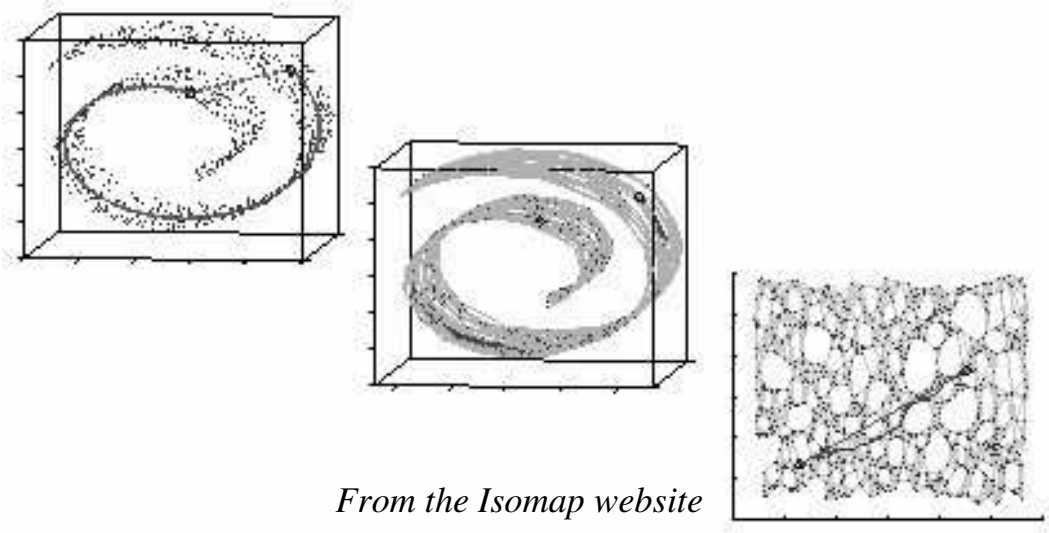
Global Coordination of Local Linear Models

Sam Roweis, Lawrence Saul
and Geoff Hinton

as narrated by Nick Roy

Manifold Learning

- Many ways to extract a low–dimensional manifold from data
  - Objective: map high– dimensional data to a low–dimensional space
  - PCA
  - multidimensional scaling
  - Isomap
  - local linear embedding



From the Isomap website

What does this paper do?

- "...synthesis of two long lines of research... ...learning the global structure of nonlinear manifolds...[and]...developing probabilistic graphical models for reasoning under uncertainty".

Factor Analysers

- What if you have noise in your observations?
  - You may want: the low–dimensional manifold *and* a density over it.
- Factor analysers:
  - Assume the data is distributed according to:

$$\boldsymbol{X} \sim N\left(\mu\ ,\ \Lambda\ I\ \Lambda^{T} + \Psi\ \right)$$

  - Task is to determine parameters of μ, Λ and Ψ

Mixture of Factor Analysers

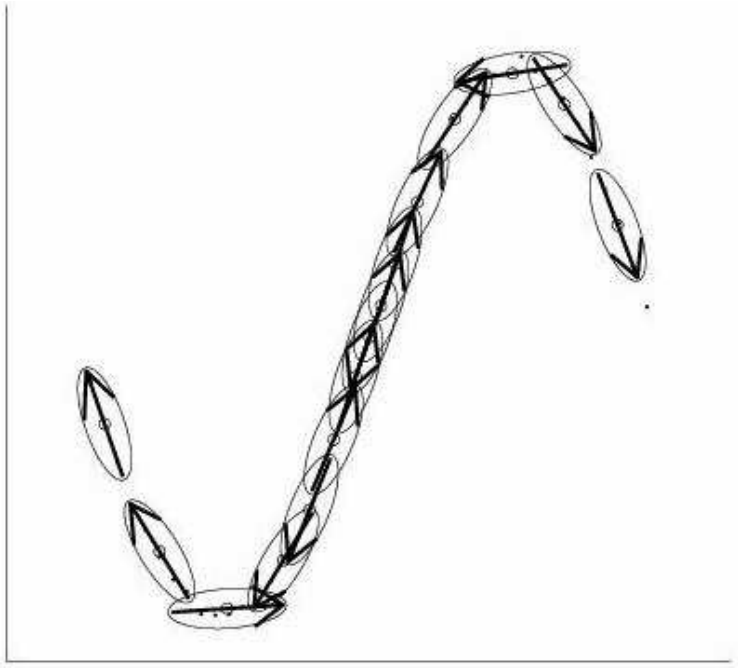
- Capture non–linear distributions using mixture of *s* analysers

$$P(\boldsymbol{x}, s, \boldsymbol{z}_s) = P(\boldsymbol{x}|s, \boldsymbol{z}_s) P(\boldsymbol{z}_s|s) P(s)$$

- Problem is now to learn μ_s, Λ_s, Ψ_s and *P(s)*
- Handled using EM

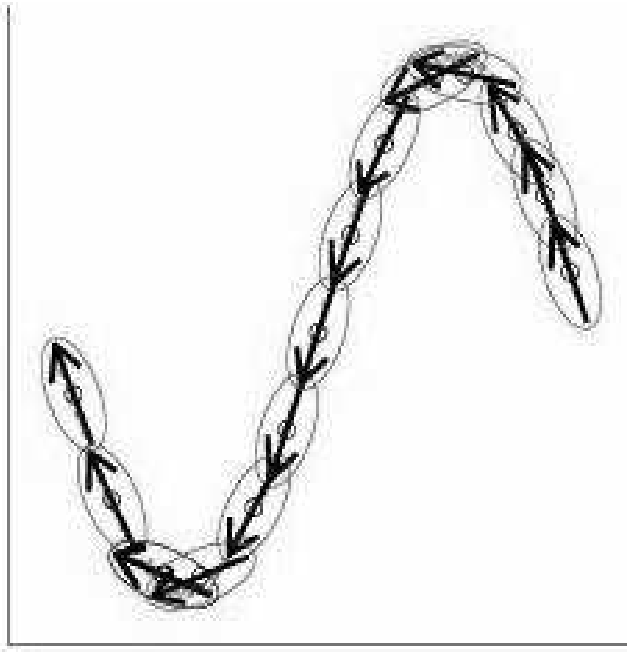
Problem

- The loading matrices Λ_s are invariant to rotations and reflections
- We can compute a density for high–dimensional co–ordinates easily from the mixture, but we cannot do the same for the manifold co–ordinates



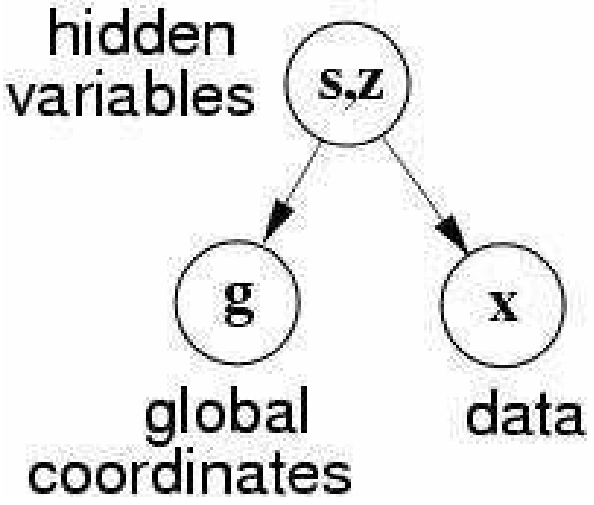
What we really want

- The factor analysers need to generate "mostly" the same global co−ordinates for the same local co−ordinates
- Allows a single set of low−dimension co−ordinates to specify "mostly" the same point in all analysers

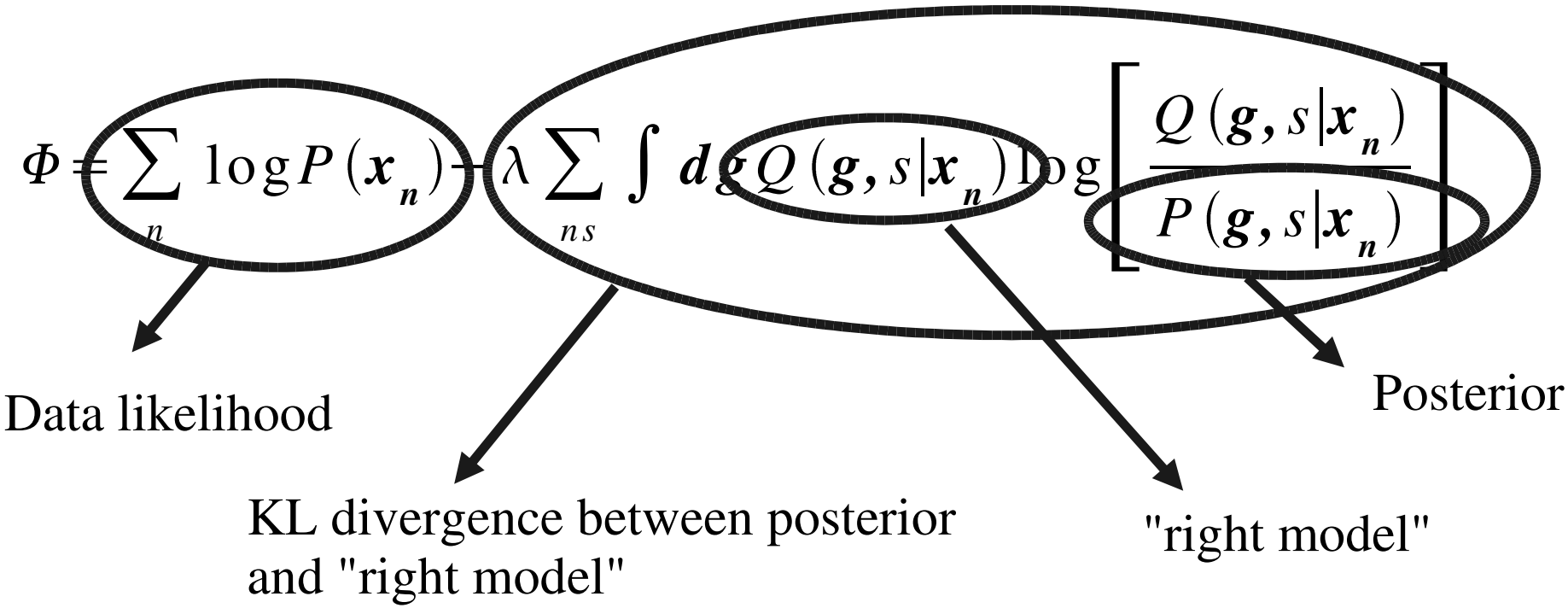


How to get all the factors to line up?

- Modify EM's objective function to penalise models that generate "multi–modal" global co–ordinates

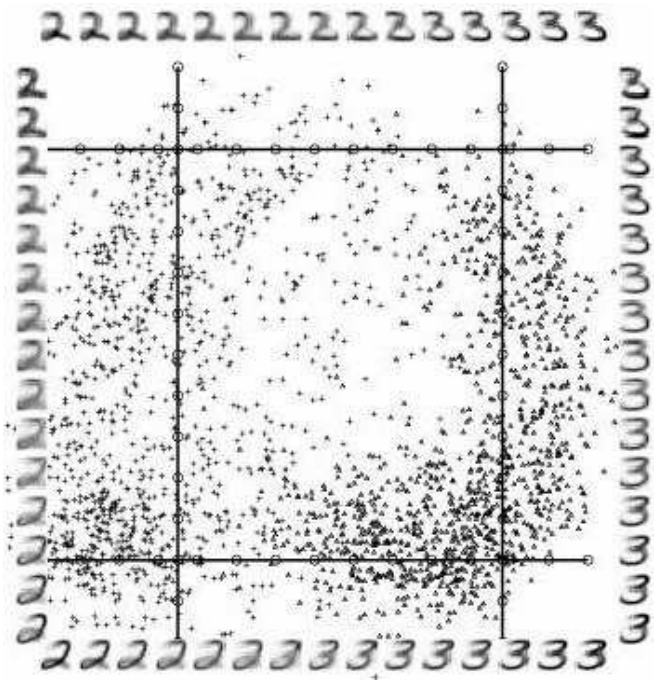


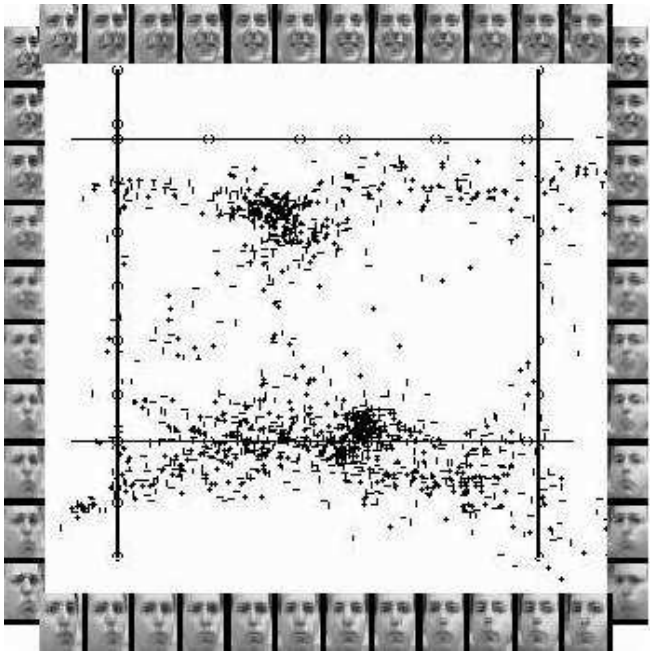
The EM objective function



$$Q(\boldsymbol{g}, s|\boldsymbol{x}_n) = Q(\boldsymbol{g}|\boldsymbol{x}_n) Q(s|\boldsymbol{x}_n)$$

Results





Hard to evaluate in the paper. Made a cool Matlab demo at NIPS.
You can download it. http://www.cs.utoronto.ca/~roweis/lle

Conclusion

- Nice way to turn Mixture of Factor Analysers into a low–dimensional manifold
- Would it work for some other local dimensionality reducer?
- Suffers from most other local dimensionality reduction methods : what happens when the manifold crosses itself?