Notes:

Latent variables - used to explain data in a simpler way

(transformation of data points into a continuous lower dimensional space)

Latent variable models - explain complex relationships between variables by simple relations between variables and an underlying unobservable - latent structure

Manifest variables - can be directly measured or observed

Latent variables - cannot be directly measured or observed

Y in latent variables, manifest or independent variables (x1,x2,x3,x4,x5) are used to explain them.

P - number of manifest variables, q- number of latent variables, q has to be less than p

Latent variable models are classifies as factor analysis, latent trait analysis, latent profile analysis, latent class analysis based on whether they are Metrical or categorical

Latent trait model - intelligence testing

Intelligence y(measured on a continuous scale)

Test: p tasks, individual scores - Xi = 1 if solves, Xi=0 if not

E(Y|X1 = x1,....,Xp=xp)

Results x = (x1,...,xp)

Model intelligence Y ~ N(u,si^2)

Estimation done by EM algorithm, E step - numerical integration, M step principle iterative methods

Estimation problems - difficult

Latent variable models - perfectly suitable for EM algorithm as Y is MCAR( Missing completely at random) but EM too slow, so hybrid approach - EM initially and special algorithms for final convergence

The linear normal factor model

P manifest variables are linearly related to the q latent variables Yt = (Y1,...,Yq) as X= u+AY+U

Linear factor analysis

Describe variation in X by variation in a latent Y plus noise, q < p

Problem : determining the smallest q for which the model is adequate, estimating factor loadings and error variances

Maximum likelihood estimation

Gaussian mixture models are a latent variable model - widely used in ML

Why latent variable models?

-Some data could be naturally observed(like patients dropping out and their missing measurements on clinical trial)

-Enables us to leverage our prior knowledge when defining a model

LVMs increase the expressive power of our model. Using a mixture of Gaussian components is much more expressive than modeling using a single component

Expectation-Maximization algorithm

Properties

Marginal likelihood increases after each EM cycle

Since Marginal likelihood is upper-bounded by its true global maximum and it increases at every step, EM must eventually converge