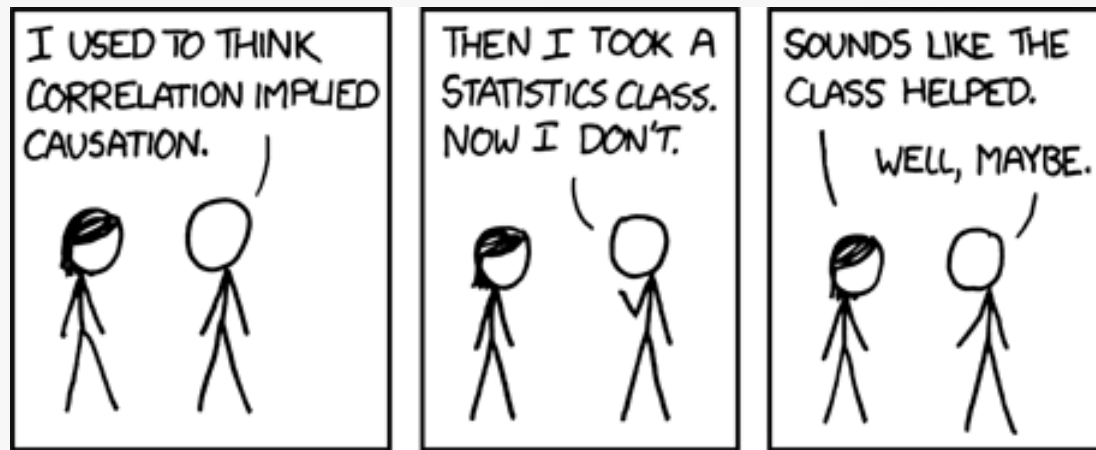


Kauzális modellek



Randall Munroe

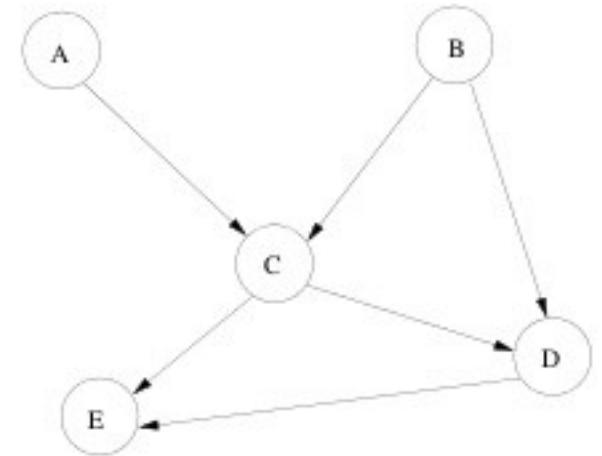
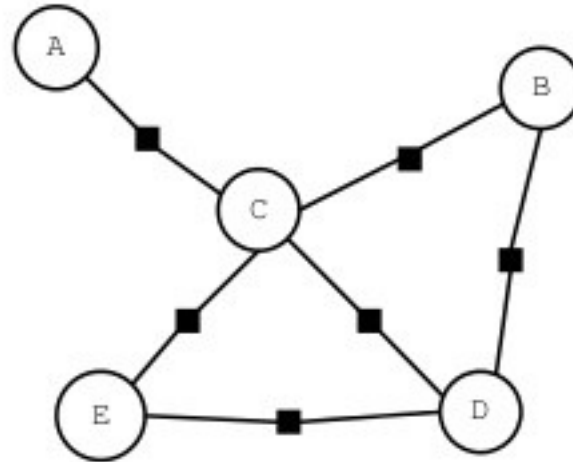
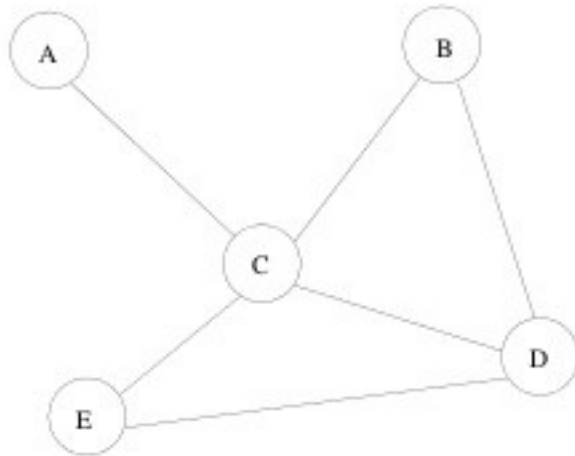
A kauzalitás reprezentációi

- Determinisztikus
 - differenciál- és differenciaegyenletek
- Sztochasztikus
 - Feltételes valószínűség $\leftrightarrow$ hipergráf
 - Irányított gráf: több ok, egy okozat $\leftrightarrow$ Bayes-háló
 - Cirkuláris kauzalitás vagy körmentes gráf (DAG)
- Dinamikus statisztikai modellek
 - Determinisztikus dinamika + zajmodell
 - Sztochasztikus differenciálegyenletek (Itô és Sztratonovics)
 - Eloszlások időbeli fejlődése (Chapman-Kolmogorov egyenlet: Fokker-Planck, Schrödinger, Ornstein-Uhlenbeck, Brown-mozgás)
- Fenomenológiai vagy mechanisztikus megközelítés
 - Ontológiai kapcsolatok – rejtett változós modellek
 - Epsztemológiai kapcsolatok - korreláció, irányított entrópia, Granger-kauzalitás, regesszió, ...

Grafikus modellek

- Minden kauzális modellre megadható
- Általában fix modell paramétereit becsüljük, vagy
- fix struktúrákat hasonlítunk össze
- Általánosan meghatározni a kauzális kapcsolatokat nagyon nehéz – Judea Pearl

Zoubin Ghahramani



Markov random field

Faktor gráf

Bayes-háló

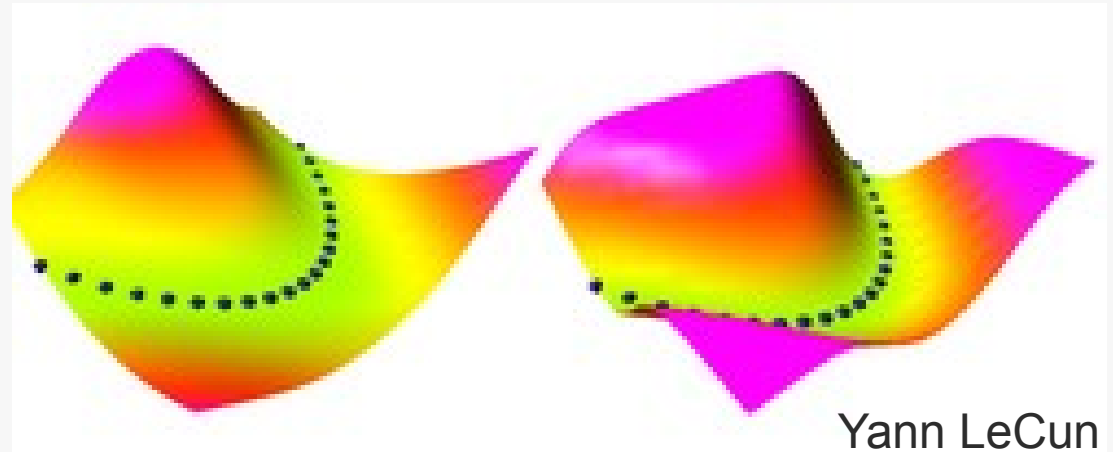
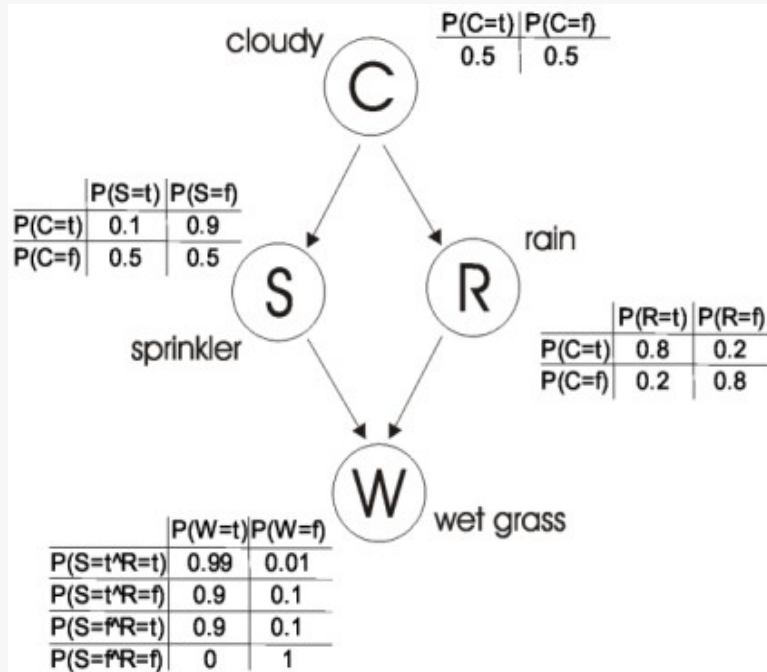
Feltételes függetlenségi viszonyok

Együttes eloszlás faktorizációja

Összefüggőségi viszonyok

Kauzális és energiaalapú modellek

Nem felügyelt tanulással



Könnyű mintagenerálás

A rejtett változók eloszlásának megtalálása (poszterior) nehezebb

Hosszadalmasabb mintagenerálás

Az energiafelületen való mozgással előállítható a poszterior

Kauzális modellek típusai

- Determinisztikus – sztochasztikus
- Diszkrét – folytonos (melyik változóról beszélünk)
- Rendezetlen adathalmaz – (időben) rendezett idősor
- Hierarchikus – hiperparaméterek adják meg a paraméterek, mint valószínűségi változók eloszlásait
- Változók közti leképezés: lineáris, bonyolultabb
- Véletlen formája: Gauss, egyenletes, multinomiális, egyéb, nemparametrikus

Eloszlások statisztikai modellekben

posterior

likelihood

prior

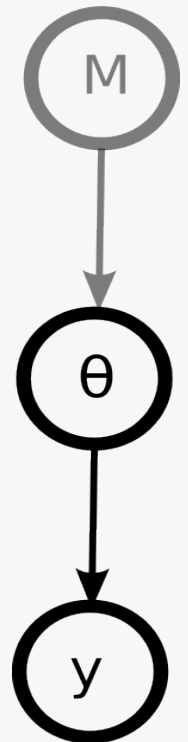
Predictive distribution:

$$p(\theta|y, M) = \frac{p(y|\theta, M) p(\theta|M)}{p(y|M)}$$

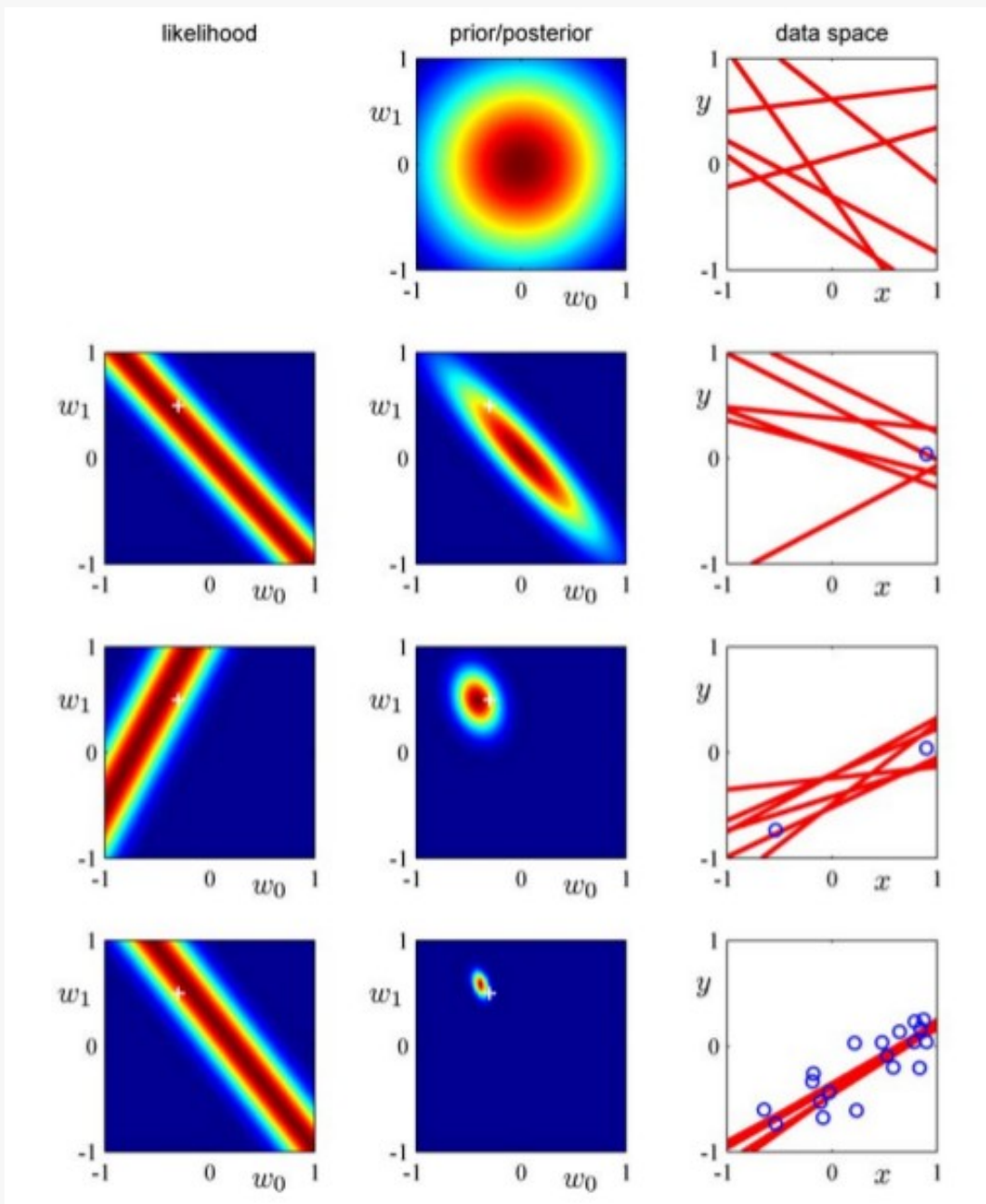
evidence
(marginal likelihood)

$$p(y'|\theta, y, M)$$

- Véletlen formája: Gauss, egyenletes, multinomiális, egyéb, nemparametrikus
- Konjugált priorok
 - A Bayes-tétel iteratív alkalmazásakor olyan prior eloszlást akarunk, ami a likelihooddal szorozva azonos alakú poszteriort ad
 - Az exponenciális eloszláscsaládoknak van konjugált priorja
 - Gauss – Gauss-Wishart (gamma, inverz gamma)
 - multinomiális – Dirichlet (binomiális-béta)

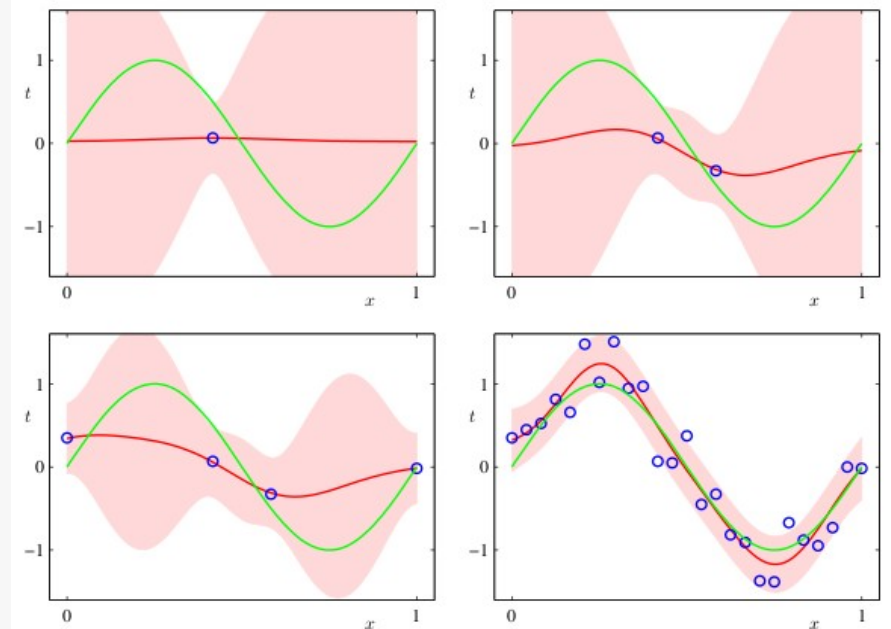


Iteratív becslés



Lineáris modell illesztése
Gauss zajjal és priorral

Harmadfokú polinom
illesztése Gauss zajjal –
prediktív eloszlás



Poszterior eloszlásbecslés

- Analitikusan sokszor nem megoldható
- Numerikus módszerek
- Két részre osztjuk a rejtett fázist – példa: szekréncipelés
- Lokális optimumot talál
- Pontbecslés
- Teljes eloszlások
 - Sampling – egzakt, könnyű megadni, de lassú
 - Közelítő inferencia, VB – nem egzakt, nehéz megadni, de gyors

Expectation Maximization

- Pontbecslés – determinisztikus paraméterek
- ML vagy MAP
- Felveszünk egy $q(x)$ segédeloszlást a rejtett változókon

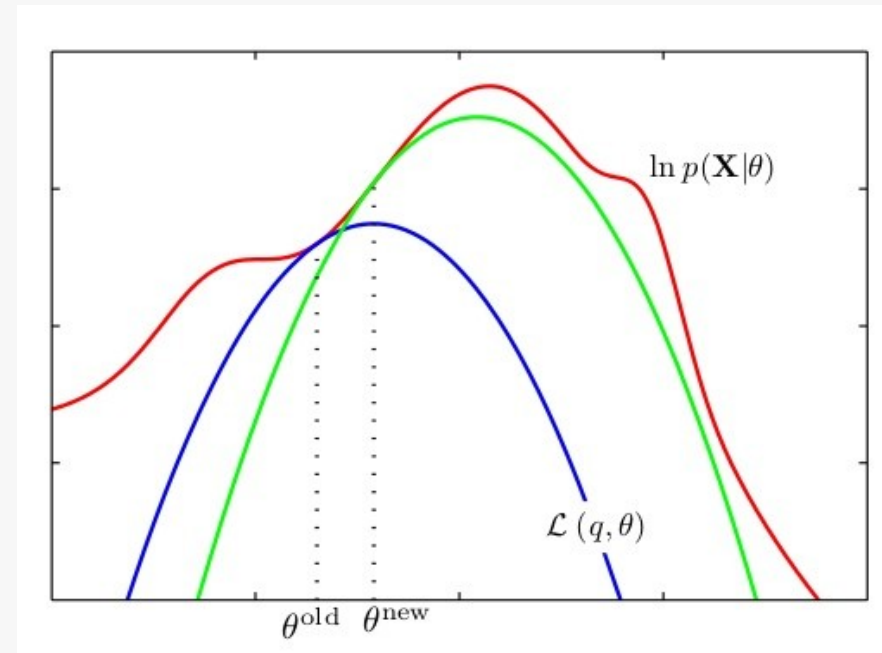
$$p(y|\theta) = L(q, \theta) + KL(q||p)$$

$$L(q, \theta) = \sum_x q(x) \ln \left\{ \frac{p(x, y|\theta)}{q(x)} \right\}$$

$$KL(q||p) = - \sum_x q(x) \ln \left\{ \frac{p(x|y, \theta)}{q(x)} \right\}$$

- E: Maximalizáljuk L-t $q(x)$ szerint a régi paraméterekkel
- M: Maximalizáljuk L-t fix $q(x)$ mellett a paraméterek szerint
- A teljes-adat-likelihood könnyebben megadható

- Ha nincs analitikus maximuma az E és M egyenleteknek, ott is numerikus közelítés kell: Gauss-Newton, Fischer scoring scheme, ...

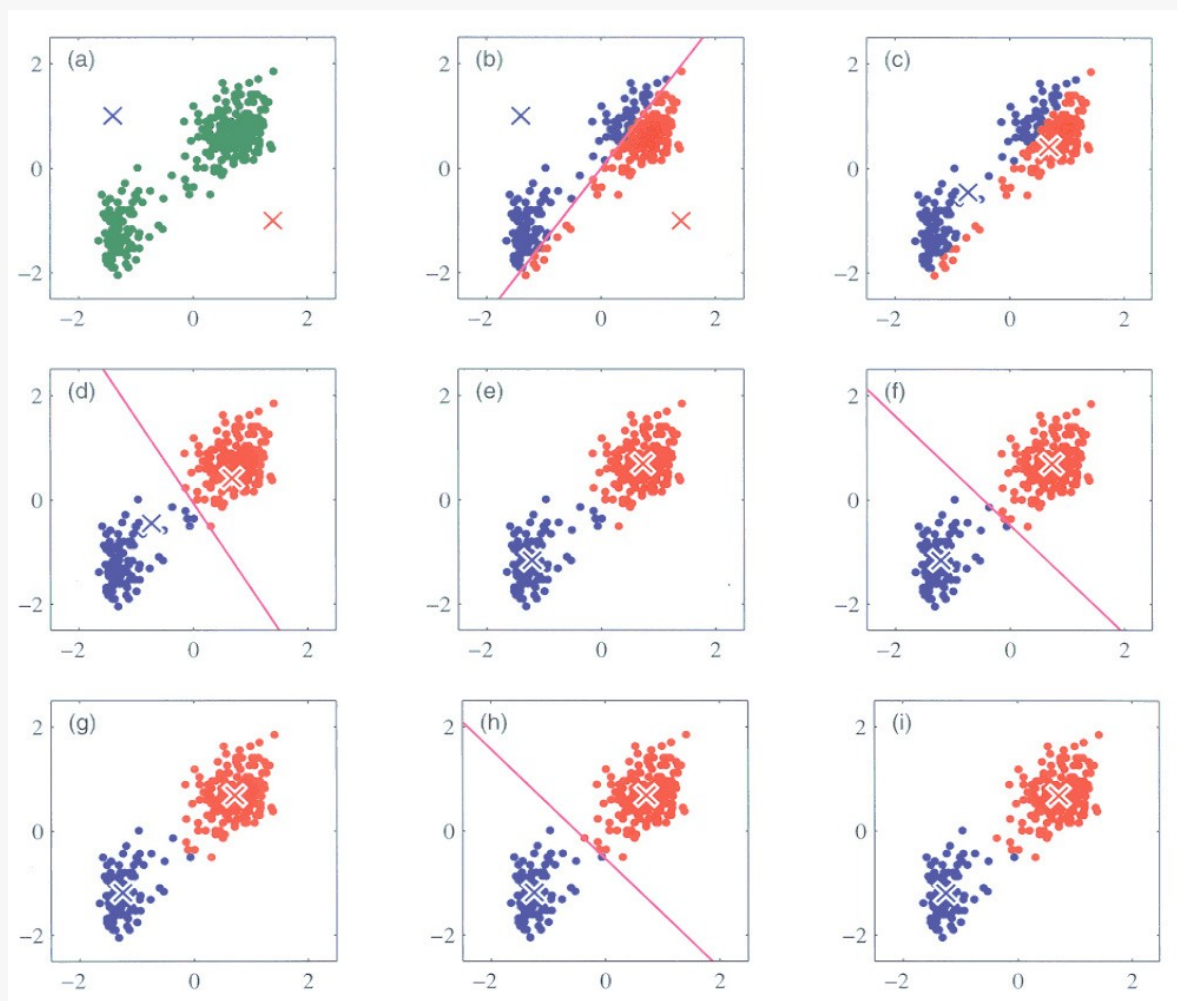


$$q(x) \leftarrow \operatorname{argmax}_q L(q, \theta) = p(x|y, \theta_{old})$$

$$\theta \leftarrow \operatorname{argmax}_\theta L(q, \theta)$$

A legegyszerűbb EM: k-means

- Kétfajta rejtett változó: klaszterközéppontok, hovatartozások
- Négyzetes távolság minimuma – max Gaussian log-likelihood egységnyi szórással



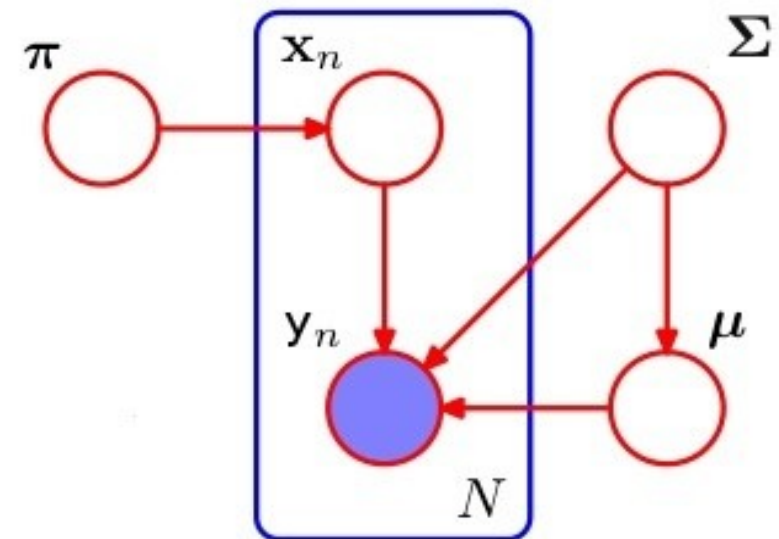
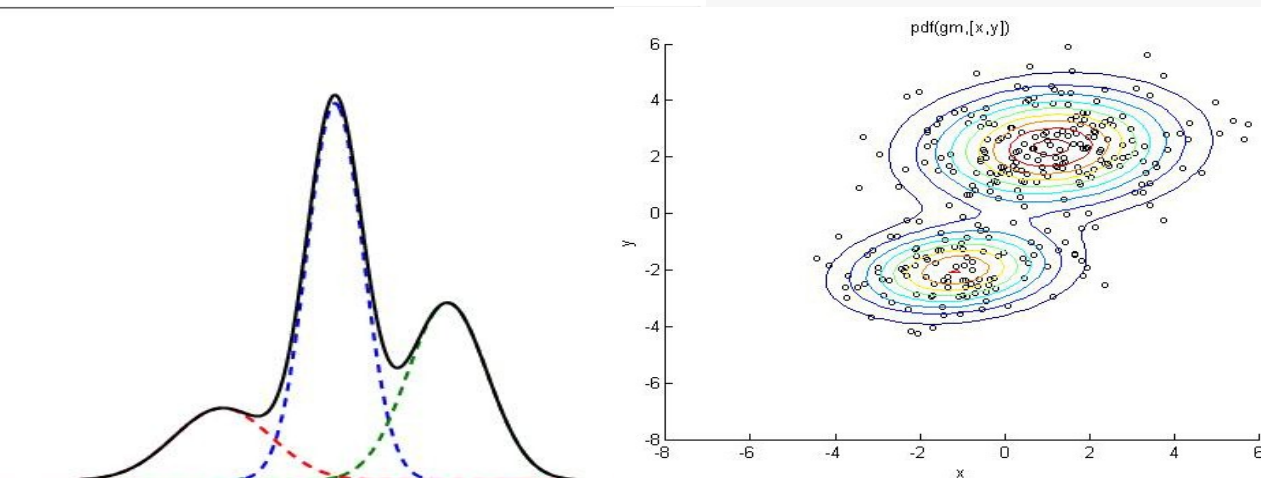
Keverékmodellek

- Eloszlások lineáris kombinációja: mixture of Gaussians
- K eloszláskomponens, N adatpont
- x - rejtett változó: melyikből generáltuk a pontot, multinomiális eloszlású k elemű vektor, egy eleme 1, a többi 0
- Multinomiális paraméterei fölött Dirichlet-priort használunk $\Lambda = \Sigma^{-1}$
- Az egyik komponensből származó Gauss-eloszlású megfigyelt változó paraméterei fölött Gauss-Wishart priort

$$p(y) = \sum_{k=1}^K \pi_k N(y | \mu_k, \Sigma_k) \quad p(\mu, \Lambda) = \prod_{k=1}^K N(\mu_k | m_0, (\beta_0 \Lambda_k)^{-1}) W(\Lambda_k | W_0, \nu_0)$$

$$p(x) = \prod_{k=1}^K \pi_k^{x_k}$$

$$p(\pi) = \text{Dir}(\pi | \alpha_0)$$



EM Gaussian mixture modellre

- E: frissítsük a rejtett változó feletti poszteriórt a paraméterek aktuális értékével

$$\gamma(x_k) = p(x_k = 1 | y) = \frac{\pi_k N(y | \mu_k, \Sigma_k)}{\sum_{j=1}^K \pi_j N(y | \mu_j, \Sigma_j)}$$

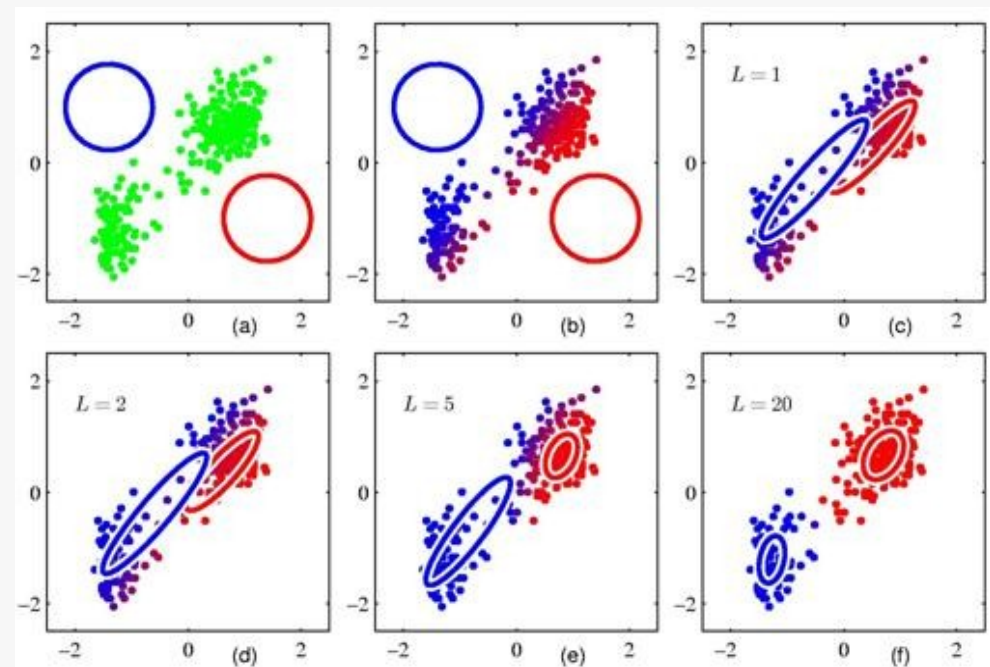
$$N_k = \sum_{n=1}^N \gamma(x_k)$$

- M: frissítsük a paraméterek értékét

$$\mu_k \leftarrow \frac{1}{N_k} \sum_{n=1}^N \gamma(x_k) y$$

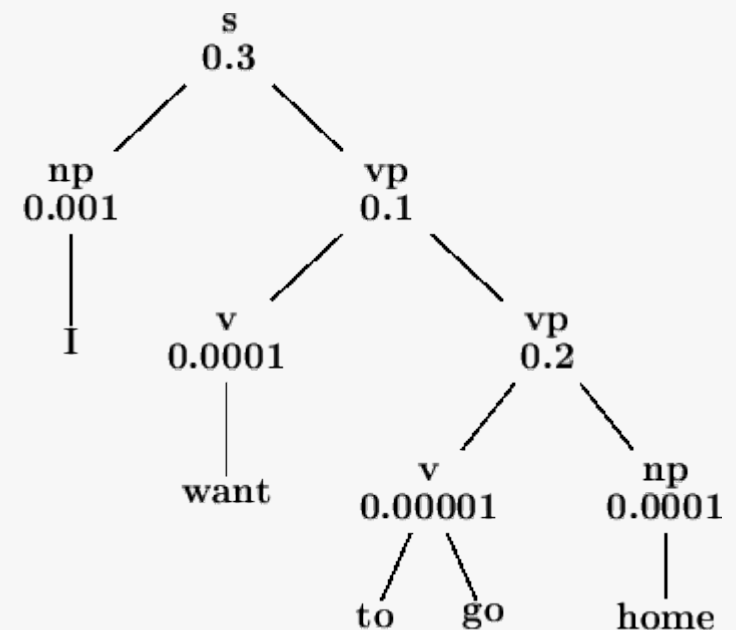
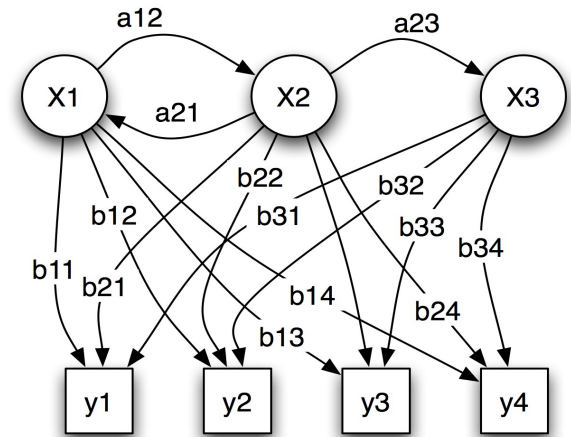
$$\Sigma_k \leftarrow \frac{1}{N_k} \sum_{n=1}^N \gamma(x_k) (y - \mu_k)(y - \mu_k)^T$$

$$\pi_k \leftarrow \frac{N_k}{N}$$



EM speciális esetei

- Hidden markov models
 - Lineáris, Gauss zaj, diszkrét változók, diszkrét idő
 - Viterbi-algoritmus: rejtett állapotváltozók
 - forward-backward (Baum-Welch): átmenetvalószínűségi paraméterek
- Kálmán-szűrő: folytonos idő
- Stochastic context-free grammar: Inside-outside algoritmus



Variational Bayes

- Poszterior közelítő becslése
- A paramétereket is valószínűségi változók, együtt kezeljük az állapotváltozókkal

$$z \leftarrow \{x, \theta\}$$

$$L(q) = \sum_z q(z) \ln \left\{ \frac{p(y, z)}{q(z)} \right\}$$

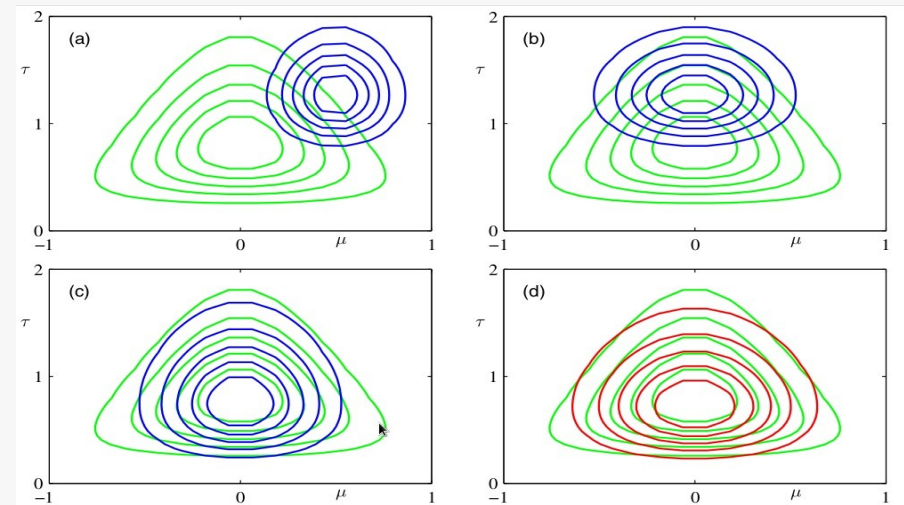
$$p(y) = L(q) + KL(q \| p)$$

$$KL(q \| p) = - \sum_z q(z) \ln \left\{ \frac{p(z|y)}{q(z)} \right\}$$

- Faktorizált eloszlást használunk közelítésre

$$q(z) = \prod_{i=1}^M q_i(z_i)$$

- A faktorok definiálhatók egyenként a rejtett változók felett, vagy csoportosítva



- L-t egyenként maximalizáljuk a faktoreloszlások szerint, a többit mindig fixen tartva

- Variációkalkulus segítségével adhatjuk meg egy funkcionál maximumát

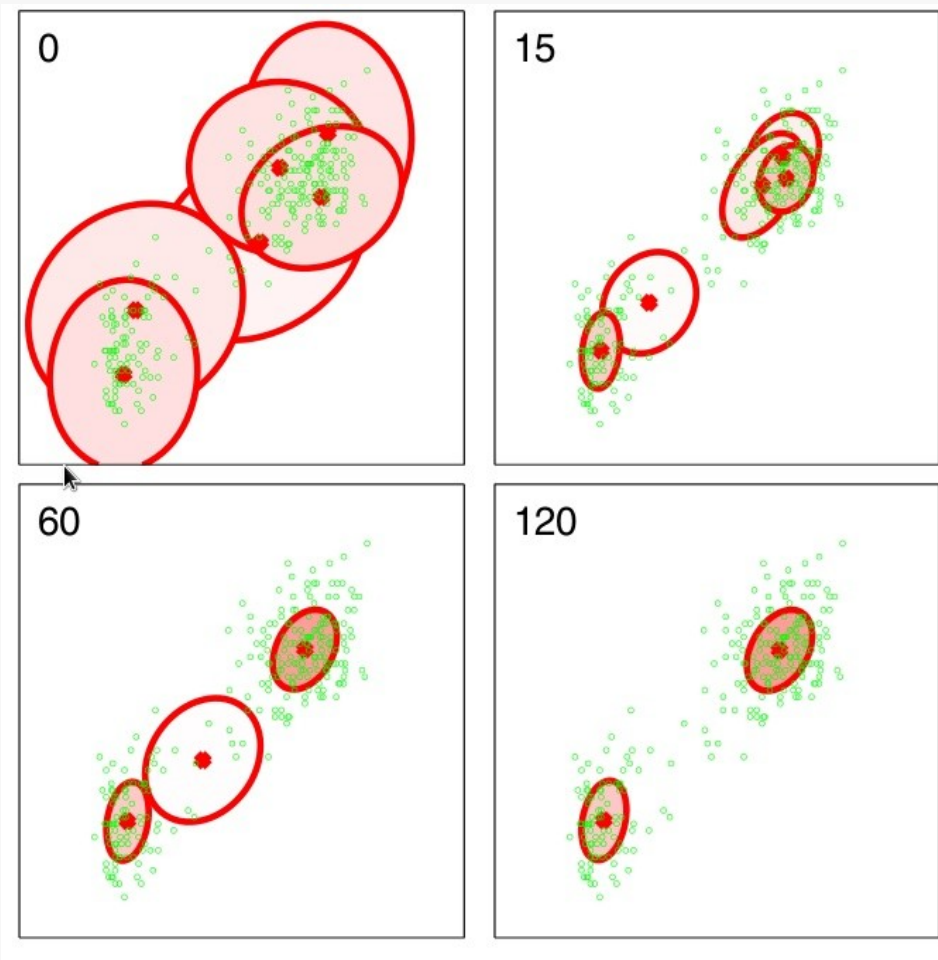
$$\ln q_j^{new}(z) = E_{q_{i \neq j}} [\ln p(y, z)] + c$$

VB for mixture of Gaussians

- Faktorok: állapotváltozók és paraméterek eloszlása

$$q(z) = q(x) q(\pi, \nu, \Lambda)$$

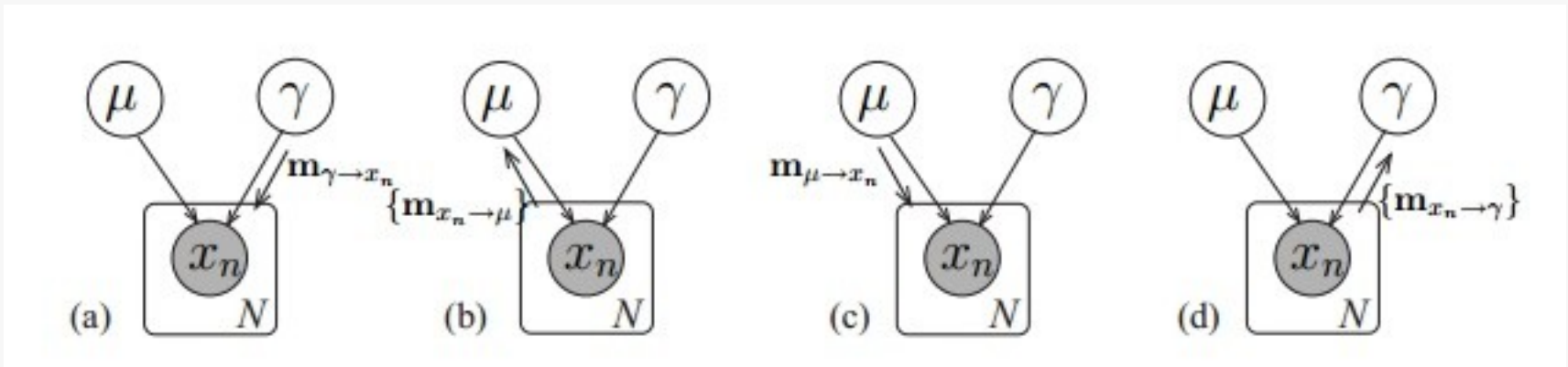
- A variációs szabállyal megadjuk a frissítési egyenleteket a q -kra
- Felváltva maximalizálunk, mit az EM-nél



$$q(x) \propto \prod_{n=1}^N \prod_{k=1}^K E[\pi_k] + \frac{1}{2} E[\ln |\Lambda_k|] - \frac{D}{2} \ln(2\pi) - \frac{1}{2} E_{\mu_k, \Lambda_k}[(y_n - \mu_k)^T \Lambda_k (y_n - \mu_k)]$$

Message passing

- Grafikus modelleken végezhető inferencia a szomszédos csúcsok közötti üzenetátadással
 - DAG-on egzakt, ciklikus gráfon közelítő
 - Hasonló, mint amikor a Boltzmann-gépet frissítjük, hogy egyensúlyba kerüljön
- Belief propagation: marginálisok
- EM: faktor gráfokon
- Variational message passing

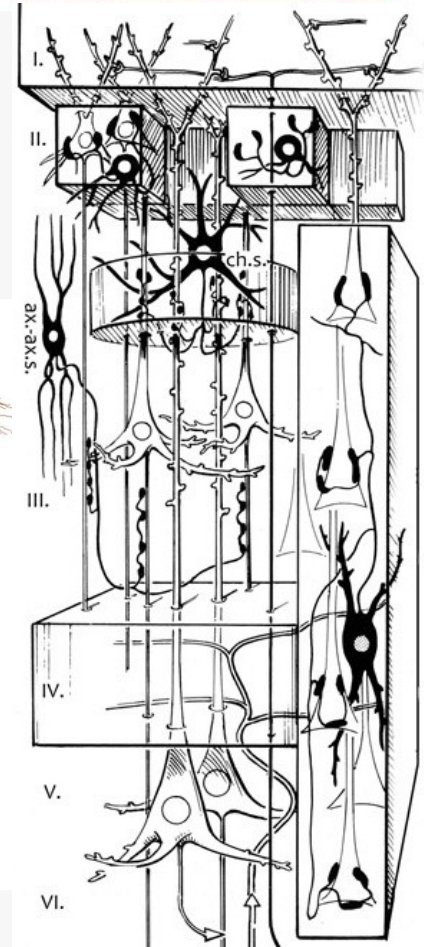
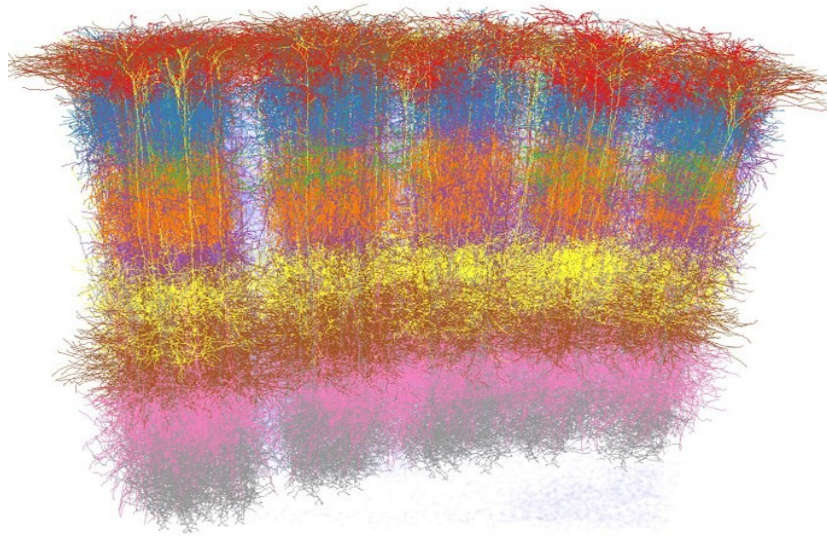
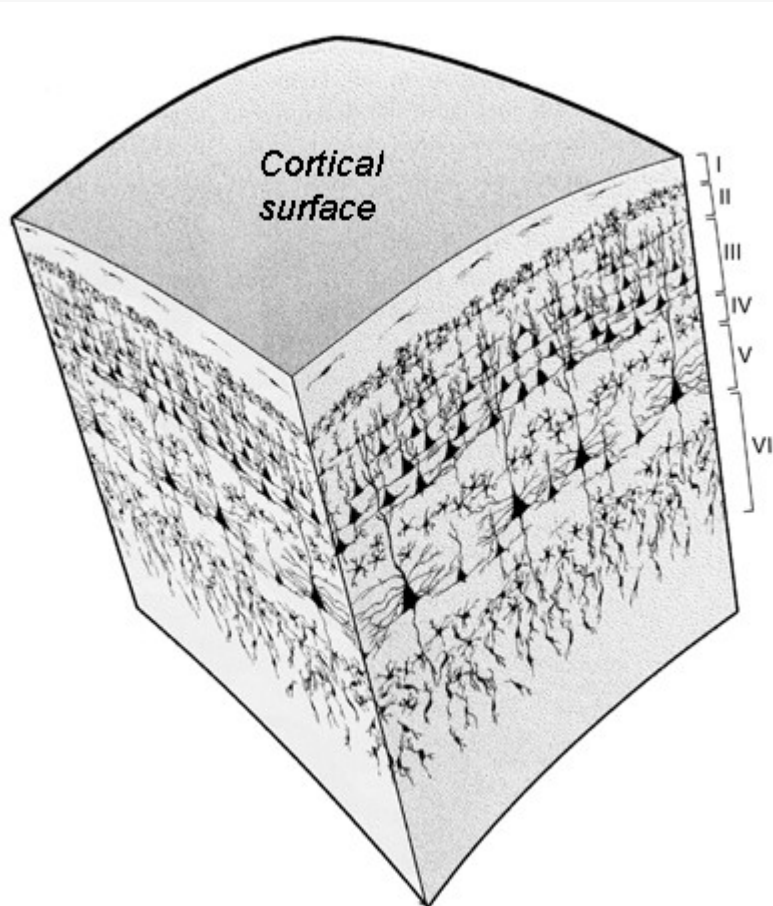
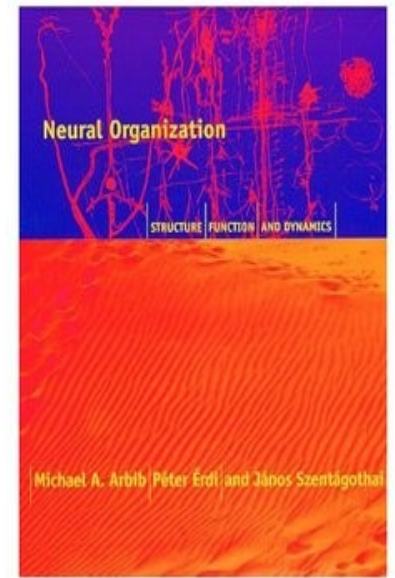


A neurális kód

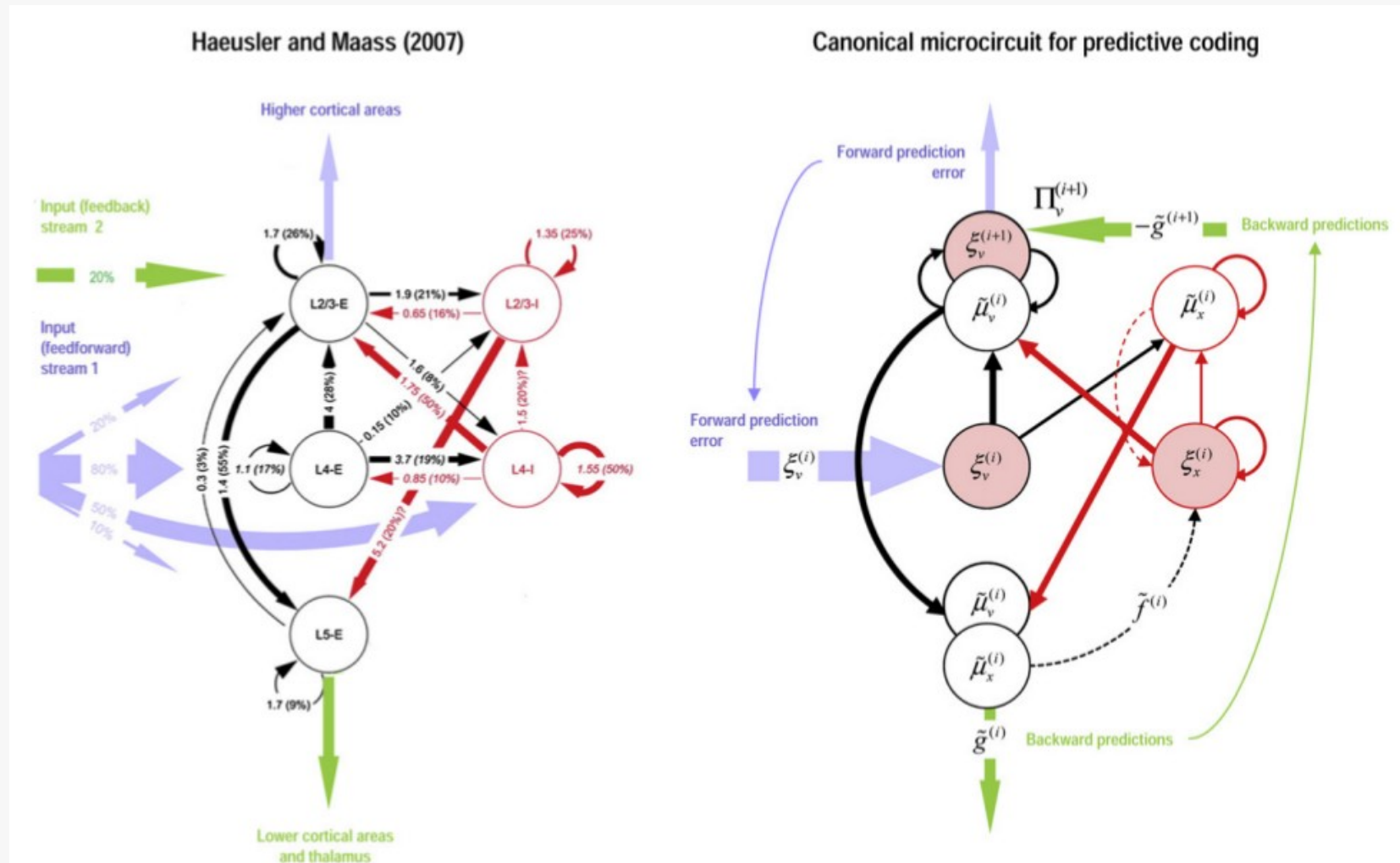
- Valószínűségi változók reprezentációja
 - Parametrikus
 - Mintaalapú
 - Predikciós hibák
- A neuronok által megvalósított számítások
 - Diszkrét spikeok és folytonos membránpotenciál
 - Lehetséges kódolóváltozók: spike timing, rate, phase, ...
 - Mi az a leképezés, amit a szinaptikus tranzmisszió megvalósít sejt- vagy populációs szinten

Az agykéreg szerkezete

- Rétegek
- Kolumnák: mikro, makro
- Szentágothai: modul



Canonical microcircuit



Karl J. Friston

- Hierarchikus modell
- Paraméterek a mély rétegekben, állapotváltozók predikciós hibája a magasabbakban
- Variational Bayes with message passing

A kérgi számítás alapelve

- Rengeteg nagyrészt inkompatibilis verzió
- Bayesiánus valószínűségre építő verziók: Bayesian brain
 - Alex Pouget, Peter Latham, Peter Dayan, Máté Lengyel, ...
- Iteratív becslésre épülő koncepciók
 - Sophie Denève: Belief propagation
 - Wolfgang Maas, Rajesh Rao: EM
 - Karl Friston: VB
 - ...