EM-алгоритм для смеси распределений в общем виде

$$p(x) = \sum_k \pi_k p_k(x|\theta) \qquad \sum_k \pi_k = 1 \qquad \pi_k \geq 0$$

$$p(x, t | \pi, \theta) = \prod_k [\pi_k p_k(x|\theta)]^{t_k} \qquad t = \{0,1\}^k$$

$$p(x|\theta, \pi) \to \max_{\theta, \pi}$$

$$\sum t_{nk} = 1$$

$$p(X, T | \pi, \theta) = \prod_n \prod_k [\pi_k p_k(x_n|\theta)]^{t_{nk}} \qquad \forall n$$

E: $q(T) = p(T | X, \pi, \theta) = \frac{1}{Z} \prod_n \prod_k [\dots]^{t_{nk}} = \prod_n q(t_n)$

$$q(t_{nk} = 1) =$$

$$= \frac{\pi_k p_k(x_n|\theta)}{\sum_j \pi_j p_j(x_n|\theta)} = \gamma_{nk}$$

EM-алгоритм для смеси распределений в общем виде

$$M: \quad \underset{\theta,\pi}{} \quad \mathbb{E}_q \log p(x, T \mid \pi, \theta) \to \max_{\theta,\pi}$$

$$\mathbb{E}_q \, t_{nk} \sum_n \sum_k \left(\log \pi_k + \log p_k(x_n \mid \theta) \right)$$

$$\sum_n \sum_k (\ldots) \cdot \underbrace{\mathbb{E}_q t_{nk}}_{\gamma_{nk}}$$

$$\pi_k = \frac{\sum_n \gamma_{nk}}{N}$$

Распределение Стьюдента и EM-алгоритм для обучения его параметров

$$\tau(x \mid \mu, \sigma^2, \nu) \propto \left(1 + \frac{1}{\nu}\left(\frac{x-\mu}{\sigma}\right)^2\right)^{-\frac{\nu+1}{2}}$$

$$t(x \mid \mu, \Sigma, \nu) \propto \left(1 + \frac{1}{\nu}(x-\mu)^T \Sigma^{-1} (x-\mu)\right)^{-\frac{\nu+D}{2}}$$

$$\tau(x \mid \mu, \Sigma, \nu) = \int_0^{+\infty} N\left(x \mid \mu, \frac{1}{z}\Sigma\right) G\left(z \mid \frac{\nu}{2}, \frac{\nu}{2}\right) dz$$

$$G(z \mid a, b) = \frac{b^a}{\Gamma(a)} z^{a-1} \exp(-bz)$$

$x_1 \dots x_N$

$$p(X \mid \mu, \Sigma, \nu) \to \max$$

$$p(X, z \mid \mu, \Sigma, \nu) = \prod_n N\left(x_n \mid \mu, \frac{1}{z_n}\Sigma\right) G\left(z_n \mid \frac{\nu}{2}, \frac{\nu}{2}\right)$$

E: $p(z \mid x, \mu, \Sigma, \nu) = q(z)$

M: $\mathbb{E}_q \log p(x, z \mid \dots)$

$\mathbb{E}_q z_n$

$\mathbb{E}_q \log z_n$

Распределение Стьюдента и EM-алгоритм для обучения его параметров

$$p(X, z \mid \dots) = \prod_n \mathcal{N}(x_n \mid \mu, \frac{1}{z_n}\Sigma)\, G(z_n \mid \frac{\nu}{2}, \frac{\nu}{2})$$

$$p(z \mid X, \dots) = \prod_n p(z_n \mid x_n, \dots)$$

$$p(z_n \mid x_n) = \frac{1}{A} \mathcal{N} \cdot G = \frac{1}{A'} \frac{1}{\left|\frac{1}{z_n}\Sigma\right|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^T \left(\frac{\Sigma}{z_n}\right)^{-1}(x-\mu)\right) \cdot$$

$$\cdot z_n^{\frac{\nu}{2}-1} \exp\left(-z_n \frac{\nu}{2}\right) = \frac{1}{A'} z_n^{\frac{\nu}{2}-1+\frac{d}{2}} \exp\left(-z_n\left(\frac{\nu}{2} + \frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right)\right)$$

$$G(z_n \mid a', b')$$

$$a' = \frac{\nu}{2} + \frac{d}{2} \qquad b' = \frac{\nu}{2} + \frac{1}{2} \dots$$

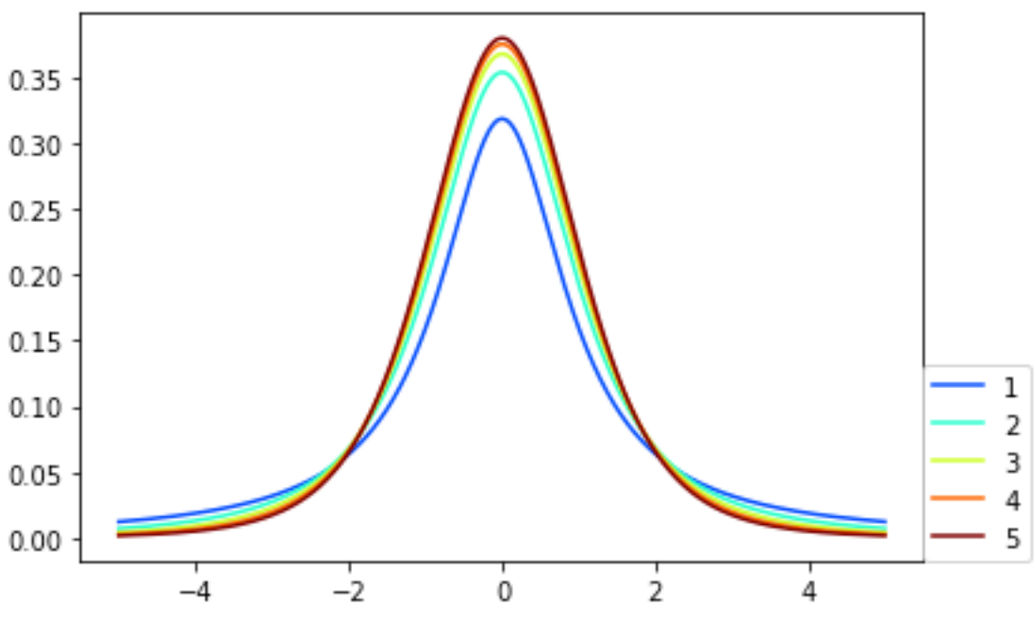
Распределение Стьюдента и EM-алгоритм для обучения его параметров

$$\mathcal{T}(\boldsymbol{x}|\boldsymbol{\mu}, \Sigma, \nu) = \frac{\Gamma\left(\frac{\nu+d}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\nu\pi}^d\sqrt{\det\Sigma}\left[1+\frac{1}{\nu}(\boldsymbol{x}-\boldsymbol{\mu})^T\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right]^{\frac{\nu+d}{2}}}, \quad \boldsymbol{x}\in\mathbb{R}^d$$

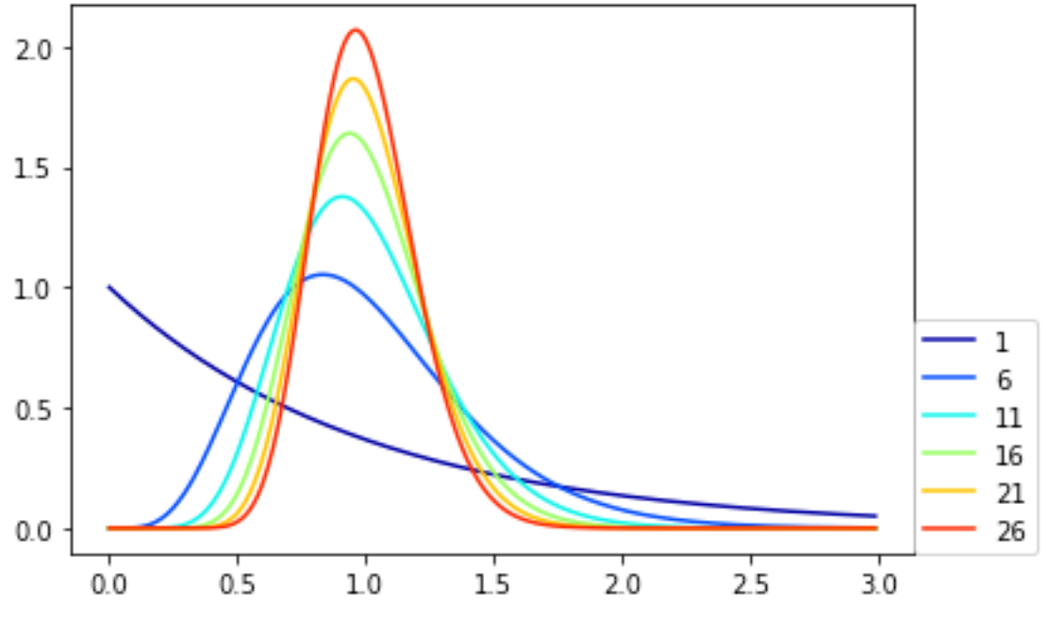
может быть представлено как

$$\int_0^{+\infty}\mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu}, z^{-1}\Sigma)\mathcal{G}(z \mid \nu/2, \nu/2)dz.$$

Распределение Стьюдента с разным ν



Гамма-распределение для весов в смеси $\mathcal{G}(z|\nu/2, \nu/2)$



Разделение смеси Стьюдентов

$$p(X,t) = \prod_n \prod_k \left[\pi_k \underbrace{p_k(x_n|\theta)}_{\tau \approx \dots}\right]^{t_{nk}}$$

$$p(z_n \mid t_{nk}=1, x_n) \propto \pi_k \cdot N \cdot G$$

$$p(X,t,z|\dots) = \prod_n \prod_k \left[\pi_k \, N\left(x_n \mid \mu_k, \frac{1}{z_n}\Sigma_k\right) G\left(z_n \mid \frac{\nu_k}{2}, \frac{\nu_k}{2}\right)\right]^{t_{nk}}$$

Е: $q(t,z) = p(tz|X) = \prod_n p(t_n, z_n | x_n)$

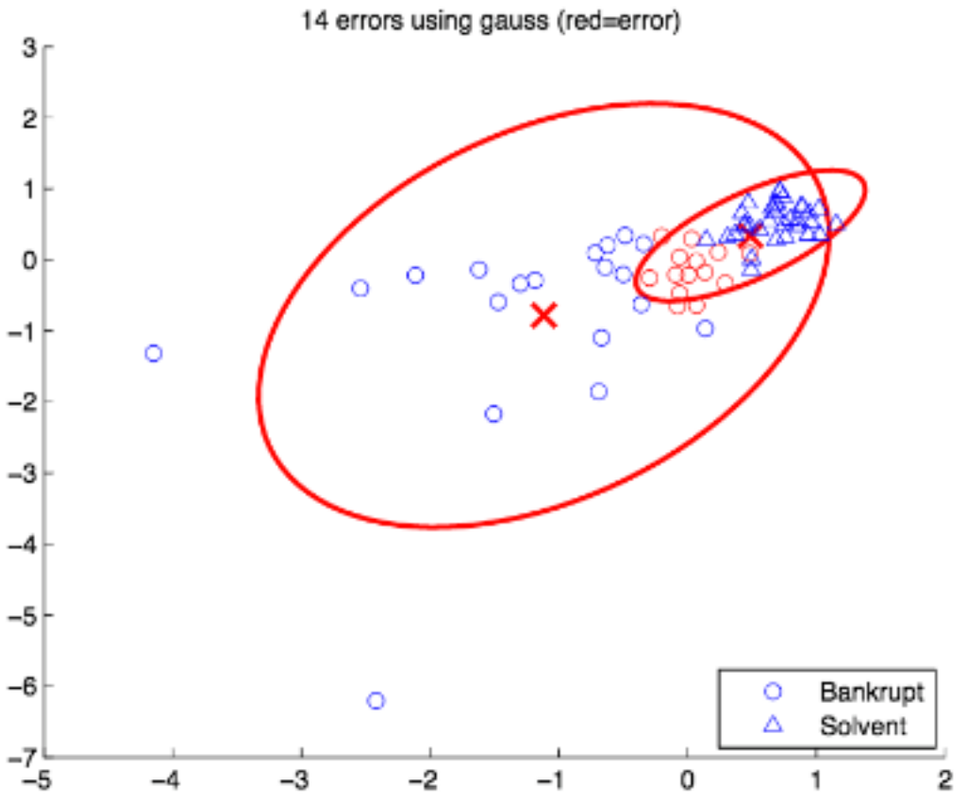
$$p(t_n, z_n | x_n) = p(z_n | t_n, x_n)\, p(t_n | x_n)$$

$$p(t_n | x_n) = \int p(t_n, z_n | x_n)\, dz_n =$$

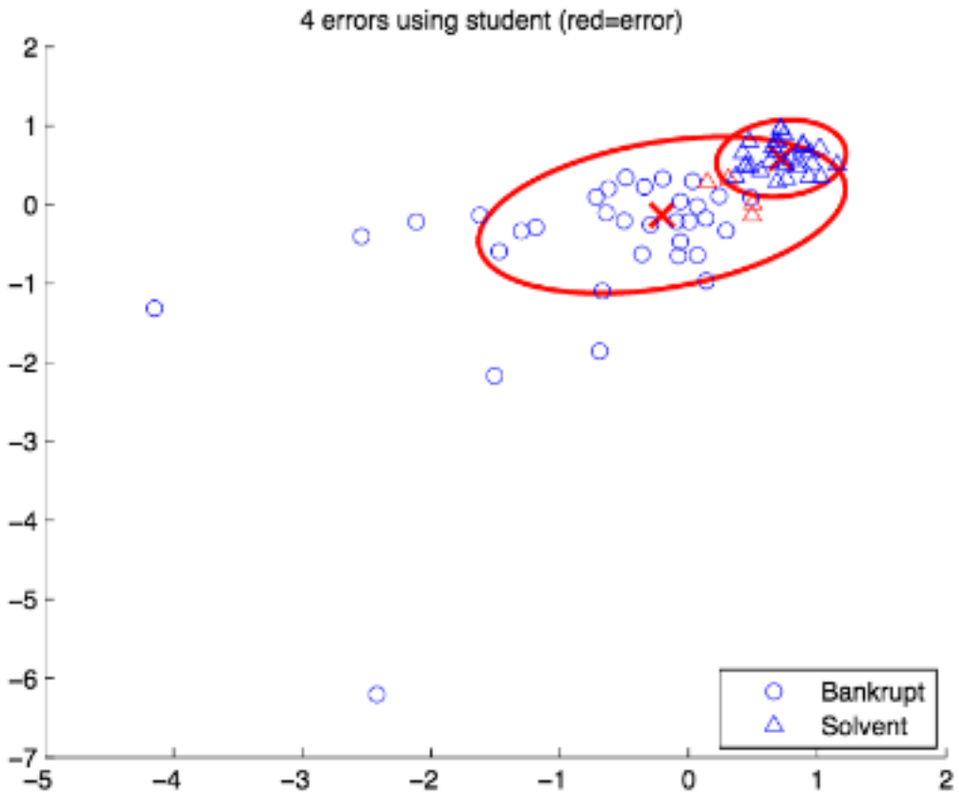
$$p(t_{nk}=1 | x_n) \propto \int p(t_{nk}=1, z_n, x_n)\, dz_n = \int \pi_k N \cdot G \, dz_n =$$

$$= \pi_k \cdot \tau(x_n|\dots) = \gamma_{nk}$$

Разделение смеси Стьюдентов



Смесь гауссиан



Смесь распределений Стьюдента