Aplicação da Abordagem Bayesiana

Modelo Gerativo para os Campos Receptivos

• Inferência Bayesiana:

→ Usada para estimação dos atributos X a partir dos observáveis Y

→ Incorpora um *modelo do mundo* e um *modelo de geração*: $Pr[X],\ Pr[Y|X]$

Regra de Bayes:

$$Pr[X,Y] = Pr[X|Y]Pr[Y] = Pr[Y|X]Pr[X]$$

Daí,

$$Pr[X|Y] = \frac{Pr[X]Pr[Y|X]}{Pr[Y]} \equiv \frac{Pr[X]Pr[Y|X]}{\sum_X Pr[X]Pr[Y|X]}$$

→ Atributos podem ser estimados maximizando-se $Pr[X|Y]$

Modelo Gerativo para os CR:

X: Campos receptivos corticais; Y: Imagens naturais

• *Modelo de Formação das Imagens*

Em uma pequena região:

$$I(\mathbf{x}) = \sum_i a_i \phi_i(\mathbf{x}) + \nu(\mathbf{x})$$

$\phi_i(\mathbf{x})$: Campo Receptivo

$\nu(\mathbf{x})$: Ruído (tudo o que não é descrito pelo modelo)

Assumindo ruído Gaussiano *i.i.d.*, com variância $1/\lambda_N$, em cada ponto:

$$Pr[\mathbf{I}|\mathbf{a}] = \Pi_{\mathbf{x}} \left\{ \sqrt{\frac{\lambda_N}{2\pi}} e^{-\frac{\lambda_N}{2}[I(\mathbf{x}) - \sum_i a_i \phi_i(\mathbf{x})]^2} \right\}$$

ou seja,

$$Pr[\mathbf{I}|\mathbf{a}] = \frac{1}{Z_N} e^{-\frac{\lambda_N}{2} \sum_{\mathbf{x}} [I(\mathbf{x}) - \sum_i a_i \phi_i(\mathbf{x})]^2}$$

com $\mathbf{I} = \{I(\mathbf{x})\},\ \mathbf{a} = \{a_i\}$

- *Modelo do Mundo*

Assumimos que os CRs proporcionam uma *representação esparsa*

$\rightarrow$ A maior parte dos a_i deve ser zero

$\rightarrow$ $Pr[\mathbf{a}]$ deve ser uma função *concentrada na origem*, com *decaimento lento*

$\rightarrow$ Escolhemos variáveis *i.i.d.* de *Cauchy*

$$Pr[\mathbf{a}] = \Pi_i \left\{ \frac{1}{\lambda_S (1 + a_i^2)} \right\}$$

ou seja,

$$Pr[\mathbf{a}] = \frac{1}{Z_S} e^{-\sum_i S(a_i)}$$

onde

$$S(a_i) = \log(1 + a_i^2)$$

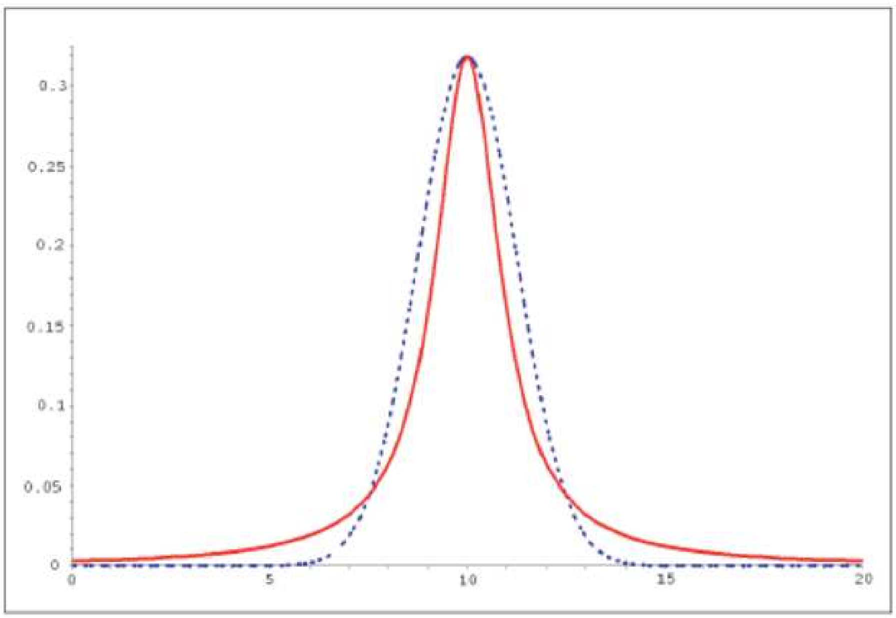
Regra de Bayes:

$$Pr[\mathbf{a}|\mathbf{I}] \propto Pr[\mathbf{I}|\mathbf{a}] Pr[\mathbf{a}]$$

Função Densidade de Cauchy:

$$f(x) = \frac{b}{\pi[(x - x_0)^2 + b^2)]}$$

Cauchy X Gaussiana



→ Pico mais estreito

→ Cauda mais larga

→ Média e variância não estão definidas

• *Estimação*

→ Precisamos obter os CRs, $\phi_i(\mathbf{x})$, e os pesos, a_i

i) Supondo conhecidos os CRs,

→ Obtemos os pesos a_i maximizando $Pr[\mathbf{a}|\mathbf{I}]$:

Empregamos *subida gradiente* sobre $\log Pr[\mathbf{a}|\mathbf{I}]$

$$\Delta a_i = \frac{\partial}{\partial a_i} \log Pr[\mathbf{a}|\mathbf{I}]$$

ou seja,

$$\Delta a_i = b_i - \sum_j C_{ij} a_j - S'(a_i)$$

onde

$$b_i = \lambda_N \sum_{\mathbf{x}} \phi_i(\mathbf{x}) I(\mathbf{x})$$

$$C_{ij} = \lambda_N \sum_{\mathbf{x}} \phi_i(\mathbf{x}) \phi_j(\mathbf{x})$$

$$S'(a_i) = \frac{d}{da_i} S(a_i)$$

→ Obtidos os a_i, podemos calcular

$$Pr[\mathbf{I}] = \sum_{\mathbf{a}} Pr[\mathbf{I}|\mathbf{a}] Pr[\mathbf{a}]$$

ii) Obtemos os CRs maximizando $Pr[\mathbf{I}]$

$\rightarrow$ Empregamos *subida gradiente* sobre $< \log Pr[\mathbf{I}|\mathbf{a}] >$

onde $< >$ indica valor médio sobre todas as imagens

Assim,

$$\Delta\phi_i(\mathbf{x}) = \frac{\partial}{\partial\phi_i(\mathbf{x})} < \log Pr[\mathbf{I}|\mathbf{a}] >$$

ou seja,

$$\Delta\phi_i(\mathbf{x}) = \lambda_N < a_i[I(\mathbf{x}) - \sum_i a_i\phi_i(\mathbf{x})] >$$

$\rightarrow$ Passos *i* e *ii* são repetidos iterativamente, partindo de CRs aleatórios

$\rightarrow$ *CRs estimados são compatíveis com os neurofisiológicos*