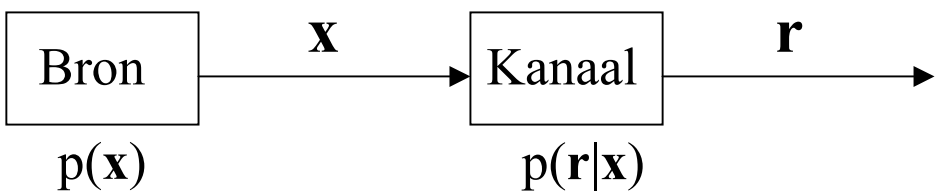
Geavanceerde Modulatie en Codering

Estimatie- en Decisietheorie

Enkele definities

Definities

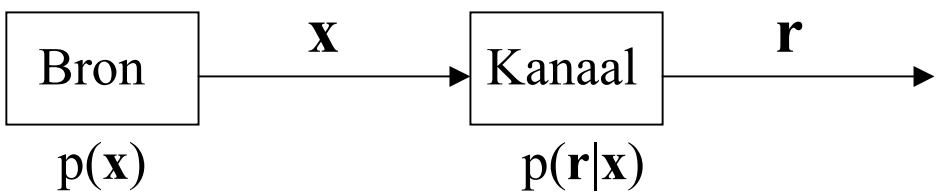


x : parametervector **r** : observatievector

p(**r**|**x**) : distributie van **r** bij gegeven **x**

p(**x**) : a priori distributie van **x**

Definities

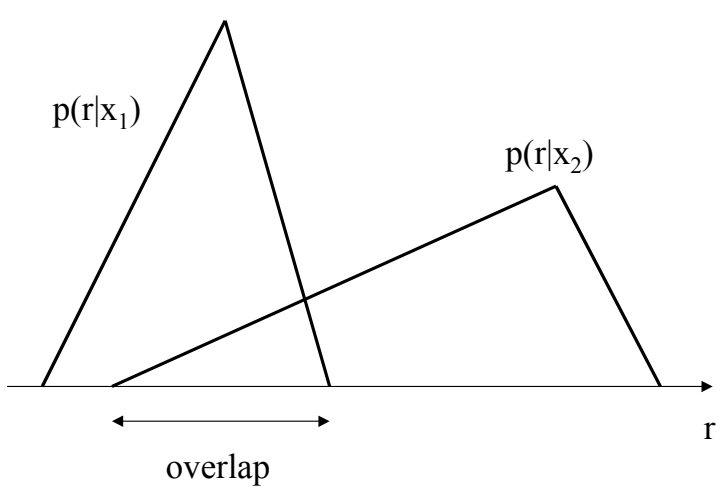


x : parametervector **r** : observatievector

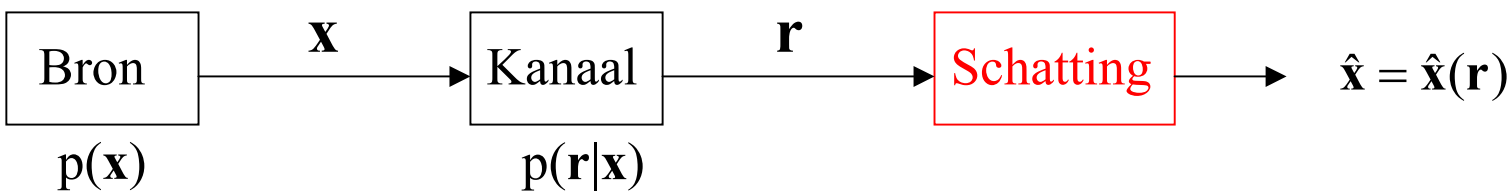
p(**r**|**x**) : distributie van **r** bij gegeven **x**

p(**x**) : a priori distributie van **x**

de exacte waarde van **x** kan meestal niet bepaald worden uit **r**



Definities

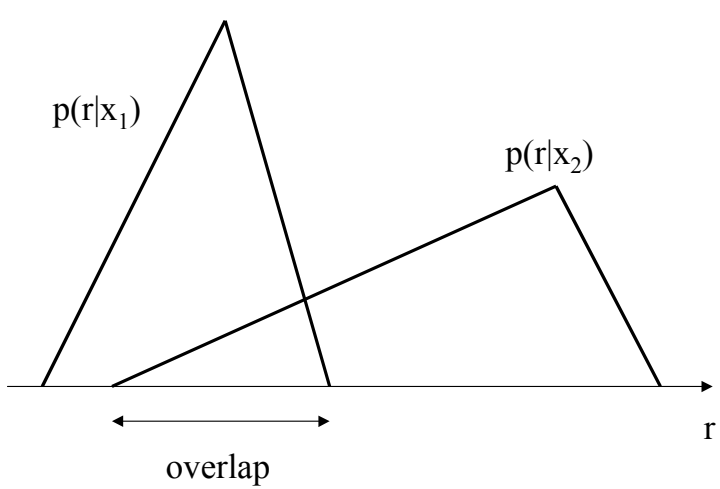


x : parametervector **r** : observatievector

$p(\mathbf{r}|\mathbf{x})$: distributie van **r** bij gegeven **x**

$p(\mathbf{x})$: a priori distributie van **x**

de exacte waarde van **x** kan meestal niet bepaald worden uit **r**



we zoeken een *schatting* $\hat{\mathbf{x}}$

$\hat{\mathbf{x}}$ is een deterministische functie van **r**

$$\hat{\mathbf{x}} = \hat{\mathbf{x}}(\mathbf{r})$$

Definities

Kostfunctie $C(\hat{\mathbf{x}}, \mathbf{x})$ vergelijkt $\mathbf{x}$ en $\hat{\mathbf{x}}$

$C(\hat{\mathbf{x}}, \mathbf{x})$ wordt minimum voor $\hat{\mathbf{x}} = \mathbf{x}$

Risico = gemiddelde van de kostfunctie

$$R = E_{\mathbf{r},\mathbf{x}}[C(\hat{\mathbf{x}}(\mathbf{r}), \mathbf{x})] = \iint C(\hat{\mathbf{x}}(\mathbf{r}), \mathbf{x}) p(\mathbf{r}, \mathbf{x}) d\mathbf{r} d\mathbf{x}$$

$E_{\mathbf{r},\mathbf{x}}[.]$: verwachtingswaarde over $\mathbf{r}$ en $\mathbf{x}$
$p(\mathbf{r},\mathbf{x})$: gezamenlijke distributie van $\mathbf{r}$ en $\mathbf{x}$

De *optimale* schatting $\hat{\mathbf{x}}(\mathbf{r})$ minimaliseert het risico R

Definities

gezamenlijke distributie $p(\mathbf{r},\mathbf{x})$ is te schrijven als

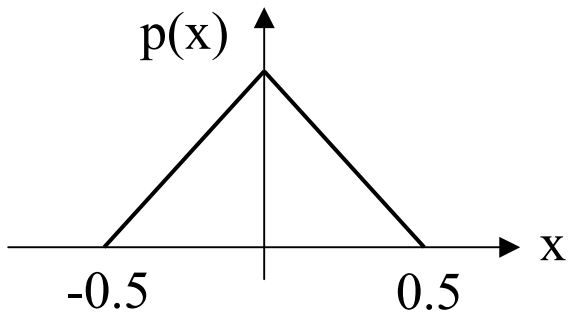
$$p(\mathbf{r},\mathbf{x}) = p(\mathbf{r}|\mathbf{x})p(\mathbf{x}) = p(\mathbf{x}|\mathbf{r})p(\mathbf{r})$$

Definities

gezamenlijke distributie p(**r**,**x**) is te schrijven als

$$p(\mathbf{r},\mathbf{x}) = p(\mathbf{r}|\mathbf{x})p(\mathbf{x}) = p(\mathbf{x}|\mathbf{r})p(\mathbf{r})$$

- p(**x**) is de *a priori* distributie van **x**, en vertegenwoordigt onze kennis van **x** wanneer we geen toegang hebben tot de observatie **r**

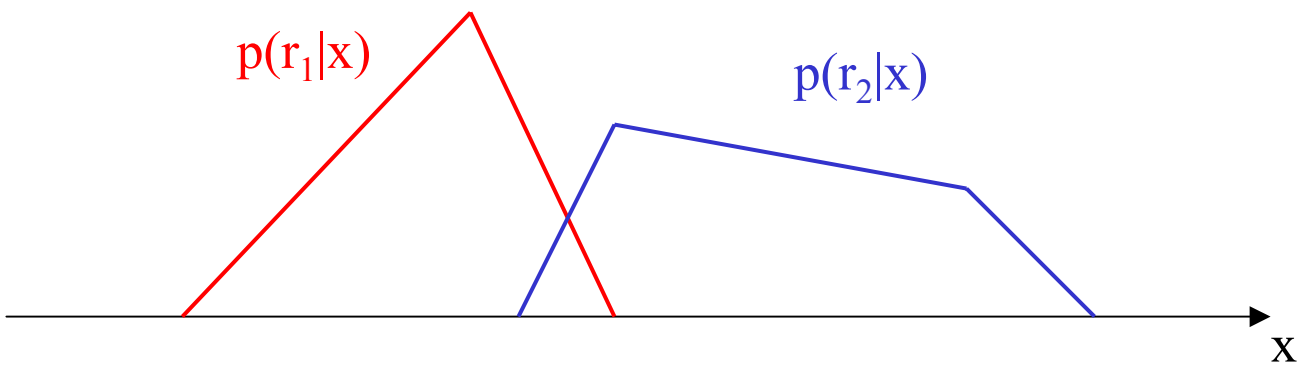


Definities

gezamenlijke distributie p(**r**,**x**) is te schrijven als

$$p(\mathbf{r},\mathbf{x}) = p(\mathbf{r}|\mathbf{x})p(\mathbf{x}) = p(\mathbf{x}|\mathbf{r})p(\mathbf{r})$$

- p(**x**) is de *a priori* distributie van **x**, en vertegenwoordigt onze kennis van **x** wanneer we geen toegang hebben tot de observatie **r**
- p(**r**|**x**), beschouwd als functie van **x**, is de *kansfunctie* (*likelihood function*) van **x**, en vertegenwoordigt onze kennis van **x** verworven via de observatie **r**



Definities

gezamenlijke distributie p(**r**,**x**) is te schrijven als

$$p(\mathbf{r},\mathbf{x}) = p(\mathbf{r}|\mathbf{x})p(\mathbf{x}) = p(\mathbf{x}|\mathbf{r})p(\mathbf{r})$$

- p(**x**) is de *a priori* distributie van **x**, en vertegenwoordigt onze kennis van **x** wanneer we geen toegang hebben tot de observatie **r**
- p(**r**|**x**), beschouwd als functie van **x**, is de *kansfunctie* (*likelihood function*) van **x**, en vertegenwoordigt onze kennis van **x** verworven via de observatie **r**
- p(**r**) is de marginale distibutie van **r** :

$$p(\mathbf{r}) = \int p(\mathbf{r},\mathbf{x})d\mathbf{x} = \int p(\mathbf{r} \mid \mathbf{x})p(\mathbf{x})d\mathbf{x}$$

p(**r**) hangt niet af van **x**, en levert dus geen informatie over **x**

Definities

gezamenlijke distributie p(**r**,**x**) is te schrijven als

$$p(\mathbf{r},\mathbf{x}) = p(\mathbf{r}|\mathbf{x})p(\mathbf{x}) = p(\mathbf{x}|\mathbf{r})p(\mathbf{r})$$

- p(**x**) is de *a priori* distributie van **x**, en vertegenwoordigt onze kennis van **x** wanneer we geen toegang hebben tot de observatie **r**
- p(**r**|**x**), beschouwd als functie van **x**, is de *kansfunctie* (*likelihood function*) van **x**, en vertegenwoordigt onze kennis van **x** verworven via de observatie **r**
- p(**r**) is de marginale distibutie van **r** :

  $$p(\mathbf{r}) = \int p(\mathbf{r},\mathbf{x})d\mathbf{x} = \int p(\mathbf{r} \mid \mathbf{x})p(\mathbf{x})d\mathbf{x}$$

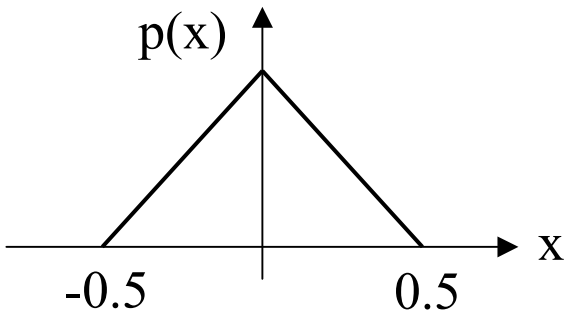
  p(**r**) hangt niet af van **x**, en levert dus geen informatie over **x**
- p(**x**|**r**) = p(**r**|**x**)p(**x**)/p(**r**) is de *a posteriori* distributie van **x**; p(**x**|**r**) vertegenwoordigt onze *totale* kennis van **x**, door a priori kennis p(**x**) en de kansfunctie p(**r**|**x**) te combineren (p(**r**) is enkel een normalisatiefactor)

Definities

Voorbeeld

$r = x + w$ $\qquad$ $w \sim N(0, \sigma^2)$

p(r|x) is Gaussiaanse distributie met gemiddelde x en variantie σ^2



a priori : x tussen -0.5 en 0.5,
hoogste waarschijnlijkheid bij x = 0

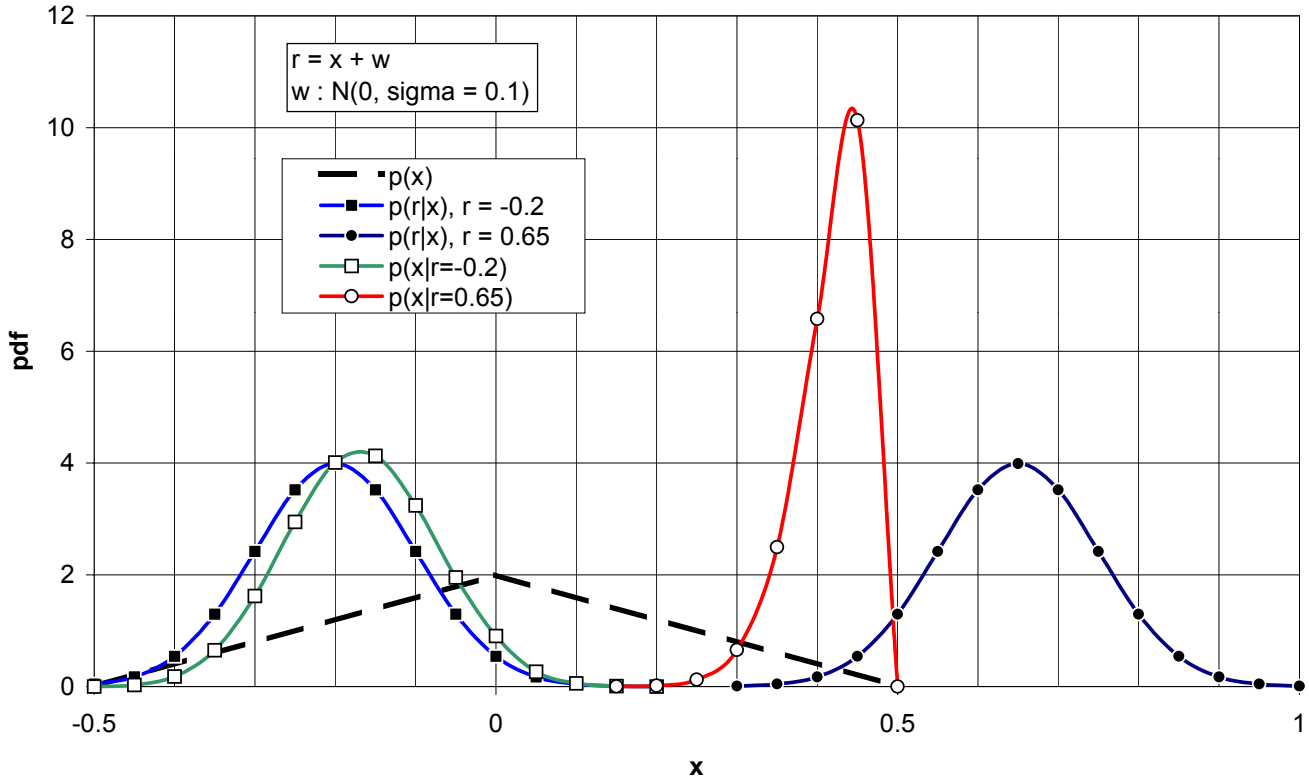
kansfunctie : x = r is de meest aannemelijke (*most likely*) waarde voor x
(bij gegeven r wordt p(r|x) maximaal voor x = r)

$$p(r \mid x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-(r-x)^2}{2\sigma^2} \right)$$

Definities

Voorbeeld (vervolg)

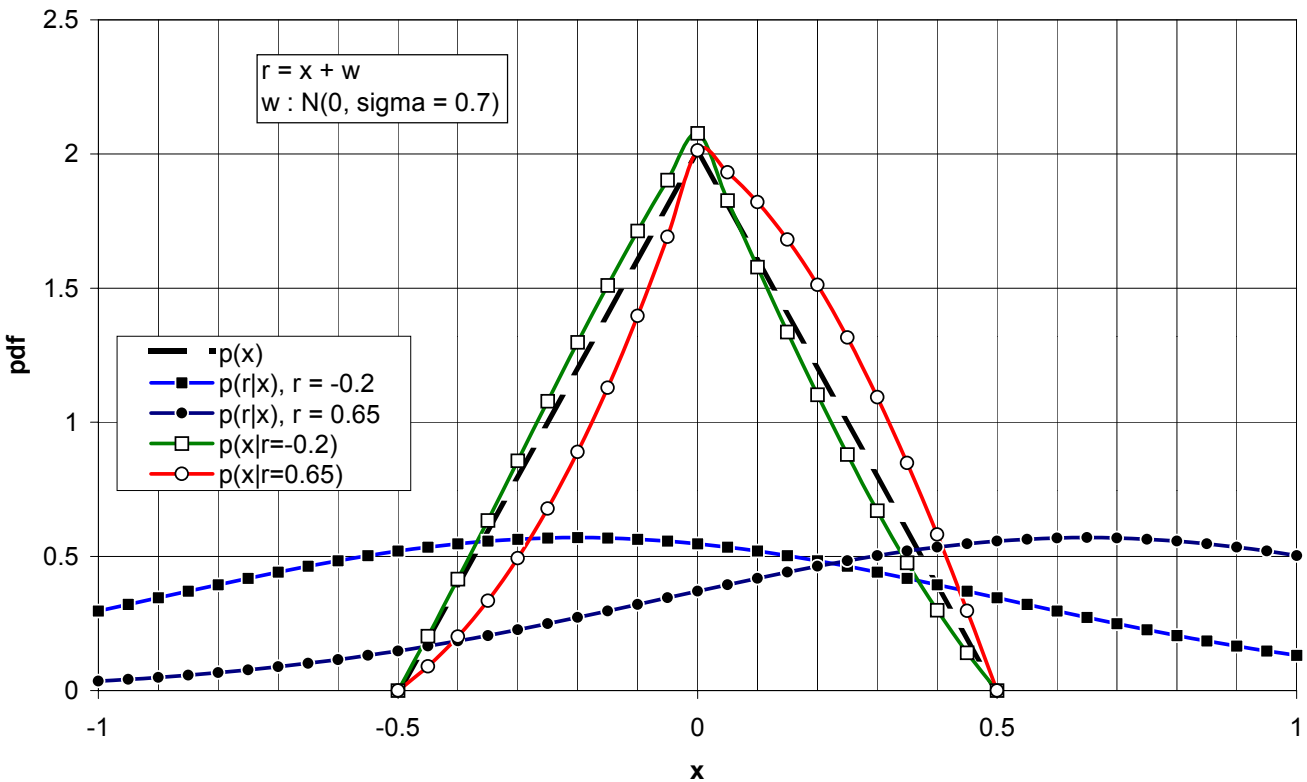
σ klein, r = -0.2 $\Rightarrow$ a posteriori distrib. $\approx$ kansfunctie
σ klein, r = 0.65 $\Rightarrow$ a posteriori distrib. verschilt sterk van kansfunctie en van a priori distrib.



Definities

Voorbeeld (vervolg)

σ groot $\Rightarrow$ a posteriori $\approx$ a priori, ongeacht waarneming r



Definities

Detectie

Schatten van een *discrete* parameter.
De schatting wordt “decisie” genoemd.

Estimatie

Schatten van een *continue* parameter.

Detectie en estimatie

Detectie

Foutprobabiliteit m.b.t. vector x

x heeft J componenten $x_1, ..., x_J$

Beschouw kostfunctie $C(\hat{\mathbf{x}}, \mathbf{x}) = \begin{cases} 0 & \hat{\mathbf{x}} = \mathbf{x} \\ 1 & \hat{\mathbf{x}} \neq \mathbf{x} \end{cases}$

Deze kostfunctie geeft aan of er een decisiefout is opgetreden, maar houdt geen rekening met het aantal foutieve componenten van **x** wanneer een decisiefout is opgetreden

Risico : $R = E[C(\hat{\mathbf{x}}, \mathbf{x})] = \Pr[\hat{\mathbf{x}} \neq \mathbf{x}]$ foutprobabiliteit m.b.t. **x**

Detectie

Foutprobabiliteit m.b.t. vector x

"Maximum a posteriori probability" (MAP) detectie van ***x***
minimaliseert foutprobabiliteit m.b.t. **x** (geen bewijs)

$$\hat{\mathbf{x}} = \arg\max_{\tilde{\mathbf{x}}} \Pr[\mathbf{x} = \tilde{\mathbf{x}} \mid \mathbf{r}] = \arg\max_{\tilde{\mathbf{x}}} \ln \Pr[\mathbf{x} = \tilde{\mathbf{x}} \mid \mathbf{r}]$$

$\Pr[\mathbf{x} = \tilde{\mathbf{x}} \mid \mathbf{r}]$: a posteriori massafunctie van **x**
(= gezamenlijke a posteriori massafunctie van $x_1, ..., x_J$)

Detectie

Gemiddelde foutprobabiliteit m.b.t. de afzonderlijke componenten van x

Beschouw kostfunctie

$$C(\hat{\mathbf{x}},\mathbf{x}) = \frac{1}{J}\sum_{j=1}^{J} I[\hat{x}_j \neq x_j] \text{ met } I[u \neq v] = \begin{cases} 0 & u = v \\ 1 & u \neq v \end{cases}$$

Deze kostfunctie is de verhouding van het aantal foutief gedetecteerde componenten van **x** tot het totaal aantal componenten van **x**

Risico :

$$R = E[C(\hat{\mathbf{x}},\mathbf{x})] = \frac{1}{J}\sum_{j=1}^{J} \Pr[\hat{x}_j \neq x_j]$$

gemiddelde foutprobabiliteit m.b.t. afzonderlijke componenten van **x**

Detectie

Gemiddelde foutprobabiliteit m.b.t. de afzonderlijke componenten van x

MAP detectie van x_j (j = 1, ..., J) minimaliseert gemiddelde foutprobabiliteit m.b.t. afzonderlijke componenten van **x** (geen bewijs)

$$\hat{x}_j = \arg\max_{\tilde{x}_j} \Pr[x_j = \tilde{x}_j \mid \mathbf{r}] = \arg\max_{\tilde{x}_j} \ln \Pr[x_j = \tilde{x}_j \mid \mathbf{r}] \qquad (j = 1, ..., J)$$

$\Pr[x_j = \tilde{x}_j \mid \mathbf{r}]$: a posteriori massafunctie van j-de component x_j

$$\Pr[x_j = \tilde{x}_j \mid \mathbf{r}] = \sum_{\mathbf{x}: x_j = \tilde{x}_j} \Pr[\mathbf{x} = \tilde{\mathbf{x}} \mid \mathbf{r}]$$

(marginale van gezamenlijke a posteriori massafunctie)

Estimatie

Beschouw kostfunctie $\mathrm{C}(\hat{\mathbf{x}}, \mathbf{x}) = |\hat{\mathbf{x}} - \mathbf{x}|^2$

Risico : $\mathrm{R} = \mathrm{E}[\mathrm{C}(\hat{\mathbf{x}}, \mathbf{x})] = \mathrm{E}[|\hat{\mathbf{x}} - \mathbf{x}|^2] = \mathrm{MSE}$

Estimatie

Beschouw kostfunctie $\quad C(\hat{\mathbf{x}}, \mathbf{x}) = |\hat{\mathbf{x}} - \mathbf{x}|^2$

Risico : $\quad R = E[C(\hat{\mathbf{x}}, \mathbf{x})] = E[|\hat{\mathbf{x}} - \mathbf{x}|^2] = \mathrm{MSE}$

Minimaliseren van MSE is vaak vrij bewerkelijk.
Alternatief : *MAP schatting* :

$$\hat{\mathbf{x}} = \arg\max_{\tilde{\mathbf{x}}} p(\tilde{\mathbf{x}} \mid \mathbf{r}) = \arg\max_{\tilde{\mathbf{x}}} \ln p(\tilde{\mathbf{x}} \mid \mathbf{r})$$

$p(\mathbf{x}|\mathbf{r})$: a posteriori densiteit van $\mathbf{x}$

Maximum-likelihood (ML) detectie/estimatie

MAP detectie van **x** kan omgevormd worden tot

$$\hat{\mathbf{x}} = \arg\max_{\tilde{\mathbf{x}}} \Pr[\mathbf{x} = \tilde{\mathbf{x}}] p(\mathbf{r} \mid \tilde{\mathbf{x}}) = \arg\max_{\tilde{\mathbf{x}}} \left(\ln \Pr[\mathbf{x} = \tilde{\mathbf{x}}] + \ln p(\mathbf{r} \mid \tilde{\mathbf{x}}) \right)$$

Indien **x** uniform verdeeld : MAP detectie wordt ML detectie

$$\hat{\mathbf{x}} = \arg\max_{\tilde{\mathbf{x}} \in X} p(\mathbf{r} \mid \tilde{\mathbf{x}}) = \arg\max_{\tilde{\mathbf{x}} \in X} \ln p(\mathbf{r} \mid \tilde{\mathbf{x}})$$

ML detectie van **x** kan ook gebruikt worden wanneer
a priori distributie van **x** onbekend is of genegeerd wordt

Maximum-likelihood (ML) detectie/estimatie

Analoge redenering voor :

1) MAP detectie van x_j :

$$\hat{x}_j = \arg\max_{\tilde{x}_j} \Pr[x_j = \tilde{x}_j] p(\mathbf{r} \mid \tilde{x}_j) = \arg\max_{\tilde{x}_j} \left(\ln \Pr[x_j = \tilde{x}_j] + \ln p(\mathbf{r} \mid \tilde{x}_j) \right)$$

Indien x_j uniform verdeeld (of a priori distributie van x_j is onbekend of wordt genegeerd) : MAP detectie wordt ML detectie

$$\hat{x}_j = \arg\max_{\tilde{x}_j} p(\mathbf{r} \mid \tilde{x}_j) = \arg\max_{\tilde{x}_j} \ln p(\mathbf{r} \mid \tilde{x}_j)$$

Maximum-likelihood (ML) detectie/estimatie

Analoge redenering voor :

1) MAP detectie van x_j :

$$\hat{x}_j = \arg\max_{\tilde{x}_j} \Pr[x_j = \tilde{x}_j] p(\mathbf{r} \mid \tilde{x}_j) = \arg\max_{\tilde{x}_j} \left(\ln \Pr[x_j = \tilde{x}_j] + \ln p(\mathbf{r} \mid \tilde{x}_j) \right)$$

Indien x_j uniform verdeeld (of a priori distributie van x_j is onbekend of wordt genegeerd) : MAP detectie wordt ML detectie

$$\hat{x}_j = \arg\max_{\tilde{x}_j} p(\mathbf{r} \mid \tilde{x}_j) = \arg\max_{\tilde{x}_j} \ln p(\mathbf{r} \mid \tilde{x}_j)$$

2) MAP estimatie van $\mathbf{x}$:

$$\hat{\mathbf{x}} = \arg\max_{\tilde{\mathbf{x}}} \Pr[\mathbf{x} = \tilde{\mathbf{x}}] p(\mathbf{r} \mid \tilde{\mathbf{x}}) = \arg\max_{\tilde{\mathbf{x}}} \left(\ln \Pr[\mathbf{x} = \tilde{\mathbf{x}}] + \ln p(\mathbf{r} \mid \tilde{\mathbf{x}}) \right)$$

Indien $\mathbf{x}$ uniform verdeeld (of a priori distributie van $\mathbf{x}$ is onbekend of wordt genegeerd) : MAP estimatie wordt ML estimatie

$$\hat{x}_j = \arg\max_{\tilde{x}_j} p(\mathbf{r} \mid \tilde{x}_j) = \arg\max_{\tilde{x}_j} \ln p(\mathbf{r} \mid \tilde{x}_j)$$

Gaussiaanse observatie

Gaussiaanse observatievector

$\mathbf{r} = \mathbf{s}(\mathbf{x}) + \mathbf{w} \qquad \mathbf{w} \sim N_c(0, 2\sigma^2 \mathbf{I}_K)$

logaritmische kansfunctie :

$$\ln p(\mathbf{r} \mid \mathbf{x}) = -K \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_k |r_k - s_k(\mathbf{x})|^2 = -K \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} |\mathbf{r} - \mathbf{s}(\mathbf{x})|^2$$

$$= -K \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \left(|\mathbf{r}|^2 - 2\,\mathrm{Re}[\mathbf{s}^H(\mathbf{x})\mathbf{r}] + |\mathbf{s}(\mathbf{x})|^2 \right)$$

termen onafhankelijk van **x** zijn irrelevant m.b.t. schatten van **x**

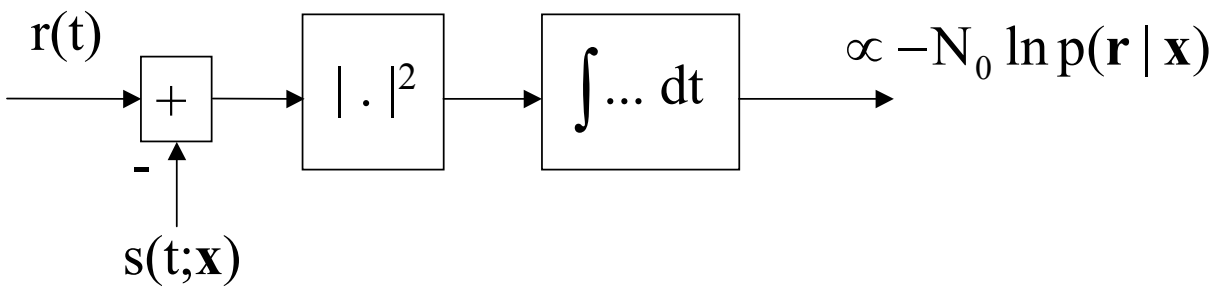
Notatie : $\ln p(\mathbf{r} \mid \mathbf{x}) \propto \frac{1}{2\sigma^2} \left(2\,\mathrm{Re}[\mathbf{s}^H(\mathbf{x})\mathbf{r}] - |\mathbf{s}(\mathbf{x})|^2 \right) \propto \frac{-1}{2\sigma^2} |\mathbf{r} - \mathbf{s}(\mathbf{x})|^2$

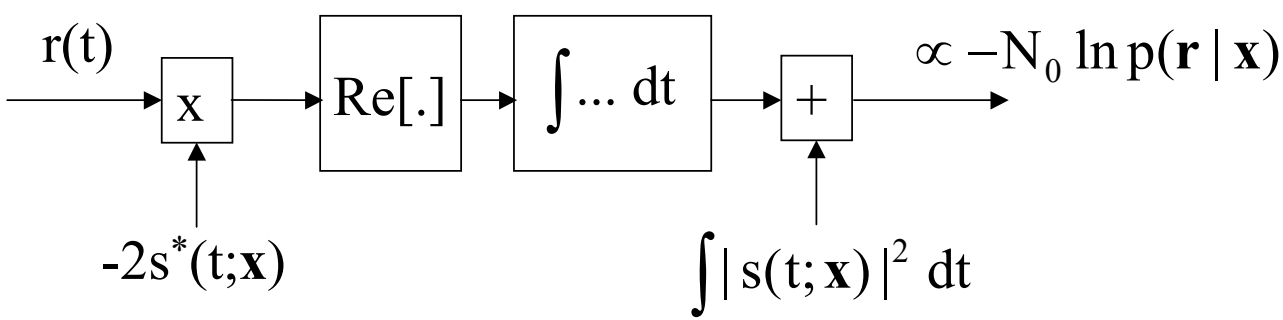
Continue-tijd Gaussiaanse observatie

$r(t) = s(t;\mathbf{x}) + w(t) \qquad w(t) \sim N_c(0, N_0\delta(u-v))$

$$N_0 \ln p(\mathbf{r} \mid \mathbf{x}) \propto -\int |r(t) - s(t;\mathbf{x})|^2 \, dt$$

$$\propto \int \left(2\,\mathrm{Re}[r(t)s^*(t;\mathbf{x})] - |s(t;\mathbf{x})|^2\right) dt$$





Continue-tijd Gaussiaanse observatie

Meerdere observaties (voorbeeld : meerdere antennes aan ontvanger)

$$r_n(t) = s_n(t;\mathbf{x}) + w_n(t) \qquad w_n(t) \sim N_c(0, N_{0,n}\delta(u-v)) \qquad (n = 0, ..., N-1)$$

$$p(\mathbf{r} \mid \mathbf{x}) = \prod_{n=0}^{N-1} p(\mathbf{r}_n \mid \mathbf{x})$$

$$\ln p(\mathbf{r} \mid \mathbf{x}) \propto \sum_{n=0}^{N-1} \frac{-1}{N_{0,n}} \int |r_n(t) - s_n(t;\mathbf{x})|^2 \, dt$$

Cramer-Rao grens

Cramer-Rao grens

$\mathbf{r} \sim N_c(\mathbf{s}(\mathbf{x}), 2\sigma^2\mathbf{I}_N)$, $\mathbf{x}$ *reële* continue parameter

1) Schatting van $\mathbf{x}$, zonder $p(\mathbf{x})$ te benutten

Stel : "echte" (unbiased) schatting : $E_{\mathbf{r}|\mathbf{x}}[\hat{\mathbf{x}}(\mathbf{r})] = \mathbf{x}$ voor alle $\mathbf{x}$

$$E_{\mathbf{r}|\mathbf{x}}\left[(\hat{x}_m(\mathbf{r}) - x_m)^2\right] \geq \left(\mathbf{J}^{-1}\right)_{m,m}$$ benchmark voor MSE

FIM : $$(\mathbf{J})_{m,n} = \frac{1}{\sigma^2}\left(\mathrm{Re}\left[\frac{\partial \mathbf{s}^H(\mathbf{x})}{\partial x_m} \cdot \frac{\partial \mathbf{s}(\mathbf{x})}{\partial x_n}\right]\right) \qquad (\mathbf{J})_{m,n} = \frac{2}{N_0}\mathrm{Re}\left[\int \frac{\partial s^*(t;\mathbf{x})}{\partial x_m} \cdot \frac{\partial s(t;\mathbf{x})}{\partial x_n} dt\right]$$

ML schatting :

$$E_{\mathbf{r}|\mathbf{x}}[\hat{\mathbf{x}}(\mathbf{r})] \to \mathbf{x} \qquad E_{\mathbf{r}|\mathbf{x}}[(\hat{\mathbf{x}}(\mathbf{r}) - \mathbf{x})(\hat{\mathbf{x}}(\mathbf{r}) - \mathbf{x})^T] \to \mathbf{J}^{-1}$$

wanneer $N \to \infty$ of $\sigma^2 \to 0$

Cramer-Rao grens

2) Schatting van **x**, met benutten van p(**x**)

$$E_{\mathbf{r},\mathbf{x}}\left[(\hat{x}_m(\mathbf{r}) - x_m)^2\right] \geq \left(\mathbf{J}_T^{-1}\right)_{m,m}$$ benchmark voor MSE

$$\mathbf{J}_T = \mathbf{J}_P + E_x[\mathbf{J}] \qquad \mathbf{J} : \text{FIM}$$

$$(\mathbf{J}_P)_{m,n} = E_{\mathbf{x}}\left[\frac{\partial \ln p(\mathbf{x})}{\partial x_m} \cdot \frac{\partial \ln p(\mathbf{x})}{\partial x_n}\right] = -E_{\mathbf{x}}\left[\frac{\partial^2 \ln p(\mathbf{x})}{\partial x_m \partial x_n}\right]$$ (bijdrage a priori informatie)

MAP schatting :

$$E_{\mathbf{r},\mathbf{x}}[(\hat{\mathbf{x}}(\mathbf{r}) - \mathbf{x})(\hat{\mathbf{x}}(\mathbf{r}) - \mathbf{x})^T] \rightarrow \mathbf{J}_T^{-1}$$

wanneer $N \rightarrow \infty$ of $\sigma^2 \rightarrow 0$

3) CRB voor complexe $\mathbf{x} = \mathbf{x}_R + j\mathbf{x}_I$: CRB toepassen op $\mathbf{x}' = (\mathbf{x}_R^T, \mathbf{x}_I^T)^T$

Voorbeelden

Decodering bij transmissie over binair symmetrisch kanaal

encoder : $\mathbf{b} \rightarrow \mathbf{c} = \chi(\mathbf{b})$ $\mathbf{b}$: k informatiebits $\mathbf{c}$: n gecodeerde bits

alle informatiebitsequenties $\mathbf{b}$ even waarschijnlijk

BSC : foutprobabiliteit p ($p < 1/2$) $\mathbf{r}$: ontvangen n-bit woord

Beschouwde decisiecriteria:

1) MAP (ML) detectie van $\mathbf{b}$

2) MAP (ML) detectie van b_i ($i = 1, ..., k$)

Decodering bij transmissie over binair symmetrisch kanaal

1) MAP (ML) detectie van **b**

$$\Pr[\mathbf{b} = \tilde{\mathbf{b}} \mid \mathbf{r}] = p(\mathbf{r} \mid \mathbf{b} = \tilde{\mathbf{b}}) \Pr[\mathbf{b} = \tilde{\mathbf{b}}] / p(\mathbf{r})$$

$$\propto p(\mathbf{r} \mid \mathbf{c} = \chi(\tilde{\mathbf{b}})) = \prod_{i=1}^{k} p(r_i \mid c_i = (\chi(\tilde{\mathbf{b}}))_i)$$

$$= (1-p)^{n-d} p^d = (1-p)^n \left(\frac{p}{1-p} \right)^d$$

$d = d_H(\mathbf{r}, \chi(\tilde{\mathbf{b}}))$ Hamming-afstand

$\hat{\mathbf{b}} = \arg\min_{\tilde{\mathbf{b}}} d_H(\mathbf{r}, \chi(\tilde{\mathbf{b}}))$ minimalisatie van Hamming afstand tussen ontvangen woord en codewoord

maximalisatie over 2^k mogelijke bitsequenties
$\Rightarrow$ in het algemeen : complexiteit $\sim 2^k$

Decodering bij transmissie over binair symmetrisch kanaal

2) MAP (ML) detectie van b_i ($i = 1, ..., k$)

$$\Pr[b_i = \tilde{b}_i \mid \mathbf{r}] = \sum_{\tilde{\mathbf{b}}: b_i = \tilde{b}_i} \Pr[\mathbf{b} = \tilde{\mathbf{b}} \mid \mathbf{r}] \propto \sum_{\tilde{\mathbf{b}}: b_i = \tilde{b}_i} p(\mathbf{r} \mid \mathbf{c} = \chi(\tilde{\mathbf{b}}))$$

- som van 2^{k-1} termen $\Rightarrow$ in het algemeen : complexiteit $\sim 2^k$
- eenvoudige (benaderde) berekening mogelijk via factorgraaf (zie later)

$$\hat{b}_i = \arg\max_{\tilde{b}_i} \Pr[b_i = \tilde{b}_i \mid \mathbf{r}] \quad i = 1, ..., k$$

k zeer eenvoudige maximalisaties
(b_i kan slechts twee waarden aannemen)

Decodering bij transmissie over AWGN kanaal

encoder + mapper : $\mathbf{b} \rightarrow \mathbf{c} \rightarrow \mathbf{a} = \chi(\mathbf{b})$

b : k informatiebits **c** : n gecodeerde bits

a : K datasymbolen, behorend tot constellatie met $M = 2^m$ punten ($K = n/m$)

alle sequenties **b** even waarschijnlijk

$\mathbf{r} = \mathbf{a} + \mathbf{w}$ $\mathbf{w} \sim N_c(0, \sigma^2 \mathbf{I}_K)$

Beschouwde decisiecriteria:

1) MAP (ML) detectie van **b**

2) MAP (ML) detectie van b_i ($i = 1, ..., k$)

Decodering bij transmissie over AWGN kanaal

1) MAP (ML) detectie van **b**

$$\Pr[\mathbf{b} = \tilde{\mathbf{b}} \mid \mathbf{r}] = \mathrm{p}(\mathbf{r} \mid \mathbf{b} = \tilde{\mathbf{b}}) \Pr[\mathbf{b} = \tilde{\mathbf{b}}] / \mathrm{p}(\mathbf{r})$$
$$\propto \mathrm{p}(\mathbf{r} \mid \mathbf{a} = \chi(\tilde{\mathbf{b}}))$$
$$\propto \exp\left(\frac{-1}{2\sigma^2} \mid \mathbf{r} - \chi(\tilde{\mathbf{b}}) \mid^2 \right)$$

$$\mathrm{d}_{\mathrm{E}}^2(\mathbf{r}, \chi(\tilde{\mathbf{b}})) = \mid \mathbf{r} - \chi(\tilde{\mathbf{b}}) \mid^2$$ kwadratische Euclidische afstand

$$\hat{\mathbf{b}} = \arg\min_{\tilde{\mathbf{b}}} \mathrm{d}_{\mathrm{E}}(\mathbf{r}, \chi(\tilde{\mathbf{b}}))$$ minimalisatie van Euclidische afstand tussen ontvangen woord en toelaatbare datasequentie

maximalisatie over 2^k mogelijke informatiebitsequenties
$\Rightarrow$ in het algemeen : complexiteit $\sim 2^k$

Decodering bij transmissie over AWGN kanaal

2) MAP (ML) detectie van b_i ($i = 1, ..., k$)

$$\Pr[b_i = \tilde{b}_i \mid \mathbf{r}] = \sum_{\tilde{\mathbf{b}}: b_i = \tilde{b}_i} \Pr[\mathbf{b} = \tilde{\mathbf{b}} \mid \mathbf{r}] \propto \sum_{\tilde{\mathbf{b}}: b_i = \tilde{b}_i} p(\mathbf{r} \mid \mathbf{a} = \chi(\tilde{\mathbf{b}}))$$

- som van 2^{k-1} termen $\Rightarrow$ in het algemeen : complexiteit $\sim 2^k$
- eenvoudige (benaderde) berekening mogelijk via factorgraaf (zie later)

$$\hat{b}_i = \arg\max_{\tilde{b}_i} \Pr[b_i = \tilde{b}_i \mid \mathbf{r}] \qquad i = 1, ..., k$$

k zeer eenvoudige maximalisaties
(b_i kan slechts twee waarden aannemen)

Fase-estimatie bij transmissie van pilootsymbolen over AWGN kanaal

$\mathbf{r} = \mathbf{a}\exp(j\theta) + \mathbf{w}$

θ : onbekende fase, uniform over $(-\pi, \pi)$
$\mathbf{a}$: gekende symboolvector (N pilootsymbolen met $|a(n)|^2 = E_s$)
$\mathbf{w} \sim N_c(0, N_0 \mathbf{I}_N)$

$$\ln p(\theta \mid \mathbf{r}) \propto \ln p(\mathbf{r} \mid \theta) \propto \frac{-1}{N_0} |\mathbf{r} - \mathbf{a}\exp(j\theta)|^2 \propto \mathrm{Re}[\mathbf{a}^H \mathbf{r} \exp(-j\theta)]$$

MAP (of ML) schatting :

$$\hat{\theta} = \arg\max_{\tilde{\theta}} p(\theta \mid \mathbf{r}) = \arg(\mathbf{a}^H \mathbf{r}) = \arg\left(\sum_{n=1}^{N} a^*(n) r(n)\right)$$

in dit voorbeeld : analytische maximalisatie van a posteriori probabiliteit

Fase-estimatie bij transmissie van pilootsymbolen over AWGN kanaal

Cramer-Rao grens

$\mathbf{r} = \mathbf{s}(\theta) + \mathbf{w}$ $\qquad$ $\mathbf{s}(\theta) = \mathbf{a}\exp(j\theta)$

FIM is scalair

$$J = \frac{2}{N_0}\left|\frac{\partial \mathbf{s}(\theta)}{\partial \theta}\right|^2 = \frac{2}{N_0}\left|j\mathbf{a}\exp(j\theta)\right|^2 = \frac{2\,|\,\mathbf{a}\,|^2}{N_0} = \frac{2}{N_0}\sum_{n=1}^{N}|\,a(n)\,|^2 = \frac{2E_s}{N_0}N$$

Cramer-Rao grens : $E[(\hat{\theta}-\theta)^2] \geq \frac{1}{N}\cdot\frac{N_0}{2E_s}$

Detectie in aanwezigheid van onbekende parameters

Detectie in aanwezigheid van onbekende parameters

encoder + mapper : $\mathbf{b} \rightarrow \mathbf{c} \rightarrow \mathbf{a} = \chi(\mathbf{b})$

observatie $\mathbf{r}$, kansfunctie $p(\mathbf{r} \mid \mathbf{a} = \chi(\mathbf{b}), \mathbf{x})$

onbekende parameter $\mathbf{x}$, onafhankelijk van $\mathbf{b}$: $p(\mathbf{x}, \mathbf{b}) = p(\mathbf{x})p(\mathbf{b})$

$$\Pr[\mathbf{b} = \tilde{\mathbf{b}} \mid \mathbf{r}] = \frac{p(\mathbf{r}, \mathbf{b} = \tilde{\mathbf{b}})}{p(\mathbf{r})} = \frac{1}{p(\mathbf{r})} \int p(\mathbf{r}, \mathbf{b} = \tilde{\mathbf{b}}, \mathbf{x}) d\mathbf{x}$$

$$= \frac{1}{p(\mathbf{r})} \int \Pr[\mathbf{b} = \tilde{\mathbf{b}} \mid \mathbf{r}, \mathbf{x}] p(\mathbf{r}, \mathbf{x}) d\mathbf{x}$$

$$= \int \Pr[\mathbf{b} = \tilde{\mathbf{b}} \mid \mathbf{r}, \mathbf{x}] p(\mathbf{x} \mid \mathbf{r}) d\mathbf{x}$$

Detectie in aanwezigheid van onbekende parameters

$$\Pr[\mathbf{b} = \tilde{\mathbf{b}} \mid \mathbf{r}] = \int \Pr[\mathbf{b} = \tilde{\mathbf{b}} \mid \mathbf{r}, \mathbf{x}] p(\mathbf{x} \mid \mathbf{r}) d\mathbf{x}$$

1) Stel waarde van $\mathbf{x}$ a priori gekend door ontvanger : $\mathbf{x} = \mathbf{x}_0$
$\Rightarrow p(\mathbf{x}|\mathbf{r}) = \delta(\mathbf{x}-\mathbf{x}_0)$

$$\Pr[\mathbf{b} = \tilde{\mathbf{b}} \mid \mathbf{r}] = \Pr[\mathbf{b} = \tilde{\mathbf{b}} \mid \mathbf{r}, \mathbf{x}_0]$$

2) Waarde van $\mathbf{x}$ niet gekend aan ontvanger :
onderstel $p(\mathbf{x}|\mathbf{r})$ vertoont scherp maximum over $\mathbf{x}$

$$\Pr[\mathbf{b} = \tilde{\mathbf{b}} \mid \mathbf{r}] \approx \Pr[\mathbf{b} = \tilde{\mathbf{b}} \mid \mathbf{r}, \hat{\mathbf{x}}]$$

$$\hat{\mathbf{x}} = \arg\max_{\tilde{\mathbf{x}}} p(\tilde{\mathbf{x}} \mid \mathbf{r})$$

Zelfde detector als bij gekende $\mathbf{x}$, maar gebruik schatting van $\mathbf{x}$ i.p.v. de correcte waarde

Detectie in aanwezigheid van onbekende parameters

$$\Pr[\mathbf{b}=\tilde{\mathbf{b}}\,|\,\mathbf{r},\mathbf{x}] = \frac{p(\mathbf{r},\mathbf{x},\mathbf{b}=\tilde{\mathbf{b}})}{p(\mathbf{r},\mathbf{x})} = \frac{p(\mathbf{r}\,|\,\mathbf{a}=\chi(\tilde{\mathbf{b}}),\mathbf{x})p(\mathbf{x})\Pr[\mathbf{b}=\tilde{\mathbf{b}}]}{p(\mathbf{r},\mathbf{x})}$$

$$\propto p(\mathbf{r}\,|\,\mathbf{a}=\chi(\tilde{\mathbf{b}}),\mathbf{x})\Pr[\mathbf{b}=\tilde{\mathbf{b}}]$$

Voorbeeld :

$\mathbf{r} = \mathbf{a}\exp(j\theta) + \mathbf{w}$

θ : onbekende fase, uniform over $(-\pi, \pi)$
$\mathbf{a} = \chi(\mathbf{b})$, met $\mathbf{b}$ een onbekende informatiebitsequentie
$\mathbf{w} \sim N_c(0,N_0\mathbf{I}_N)$

$$-N_0 \ln p(\mathbf{r}\,|\,\mathbf{a}=\chi(\mathbf{b}),\theta) \propto |\mathbf{r}-\chi(\mathbf{b})\exp(j\theta)|^2 = |\mathbf{r}\exp(-j\theta)-\chi(\mathbf{b})|^2$$

$$\propto |\chi(\mathbf{b})|^2 - 2\,\mathrm{Re}[(\chi(\mathbf{b}))^H\mathbf{r}\exp(-j\theta)]$$

Estimatie in aanwezigheid van onbekende symbolen

Estimatie in aanwezigheid van onbekende symbolen

$$\hat{\mathbf{x}} = \arg\max_{\tilde{\mathbf{x}}} \mathrm{p}(\tilde{\mathbf{x}} \mid \mathbf{r})$$

$$\mathrm{p}(\tilde{\mathbf{x}} \mid \mathbf{r}) \propto \sum_{\tilde{\mathbf{b}}} \mathrm{p}(\tilde{\mathbf{x}}, \mathbf{r}, \mathbf{b} = \tilde{\mathbf{b}}) = \mathrm{p}(\tilde{\mathbf{x}}) \sum_{\tilde{\mathbf{b}}} \mathrm{p}(\mathbf{r} \mid \mathbf{a} = \chi(\tilde{\mathbf{b}})) \Pr[\mathbf{b} = \tilde{\mathbf{b}}]$$

b bevat k informatiebits
$\Rightarrow$ som bestaat uit 2^k termen
$\Rightarrow$ complexiteit $\sim 2^k$

Complexiteit kan omzeild worden door gebruik te maken van "expectation-maximization" (EM) algoritme

EM algoritme

- MAP schatting van **x** gebeurt door middel van iteraties
- Generatie van sequentie van schattingen :

$$\hat{\mathbf{x}}^{(i)} \qquad i = 0, 1, 2, \ldots$$

- per iteratie twee stappen : E-stap en M-stap
- E-stap (expectation)

$$\begin{aligned} Q(\tilde{\mathbf{x}}, \hat{\mathbf{x}}^{(i)}) &= E_{\mathbf{a}|\mathbf{r},\hat{\mathbf{x}}^{(i)}}[\ln p(\mathbf{r} \mid \mathbf{a}, \tilde{\mathbf{x}})] \\ &= \sum_{\tilde{\mathbf{a}}} (\ln p(\mathbf{r} \mid \tilde{\mathbf{a}}, \tilde{\mathbf{x}})) \Pr[\mathbf{a} = \tilde{\mathbf{a}} \mid \mathbf{r}, \hat{\mathbf{x}}^{(i)}] \end{aligned}$$

- M-stap (maximization)

$$\hat{\mathbf{x}}^{(i+1)} = \arg\max_{\tilde{\mathbf{x}}} (\ln p(\tilde{\mathbf{x}}) + Q(\tilde{\mathbf{x}}, \hat{\mathbf{x}}^{(i)}))$$

- schatting convergeert naar stationair punt van p(**x**|**r**), dus niet noodzakelijk naar het globaal maximum (invloed van initiële schatting)

EM algoritme

Voorbeeld

$\mathbf{r} = \mathbf{a}\exp(j\theta) + \mathbf{w}$

θ : onbekende fase, uniform over $(-\pi, \pi)$
$\mathbf{a} = \chi(\mathbf{b})$, met $\mathbf{b}$ een onbekende informatiebitsequentie; $|a(n)|^2 = E_s$ (M-PSK)
$\mathbf{w} \sim N_c(0, N_0\mathbf{I}_N)$

$$\ln p(\mathbf{r} \mid \mathbf{a}, \theta) \propto \frac{2}{N_0} \mathrm{Re}\left[\sum_{n=1}^{N} a^*(n) r(n) \exp(-j\theta)\right]$$

$$E_{\mathbf{a}|\mathbf{r},\hat{\theta}^{(i)}}\left[\ln p(\mathbf{r} \mid \mathbf{a}, \theta)\right] \propto \frac{2}{N_0} \mathrm{Re}\left[\sum_{n=1}^{N} \mu^{(i)*}(n) r(n) \exp(-j\theta)\right] \qquad \text{(E-stap)}$$

$$\mu^{(i)}(n) = E_{\mathbf{a}|\mathbf{r},\hat{\theta}^{(i)}}[a(n)] = \sum_{\alpha} \alpha \Pr[a(k) = \alpha \mid \mathbf{r}, \hat{\theta}^{(i)}]$$

a posteriori verwachtings-waarde van a(k)

$$\Pr[a(k) = \alpha \mid \mathbf{r}, \hat{\theta}^{(i)}]$$

marginale a posteriori waarschijnlijkheid

EM algoritme

Voorbeeld (vervolg)

$$\hat{\theta}^{(i+1)} = \arg\max_{\tilde{\theta}} \mathrm{Re}\left[\sum_{n=1}^{N} \mu^{(i)*}(n) r(n) \exp(-j\tilde{\theta})\right]$$

$$= \arg\left(\left[\sum_{n=1}^{N} \mu^{(i)*}(n) r(n)\right]\right) \qquad \text{(M-stap)}$$

onbekende symbolen worden vervangen door hun a posteriori verwachtingswaarden, die worden gebruikt alsof ze de correcte waarde van de datasymbolen voorstellen

EM algoritme

Voorbeeld (vervolg)

Berekening a posteriori distributie van a(n)

$$\Pr[a(n)=\alpha \mid \mathbf{r}, \tilde{\theta}^{(i)}] = \sum_{\tilde{\mathbf{a}}:\tilde{a}(n)=\alpha} \Pr[\mathbf{a}=\tilde{\mathbf{a}} \mid \mathbf{r}, \tilde{\theta}^{(i)}] \propto \sum_{\tilde{\mathbf{a}}:\tilde{a}(n)=\alpha} p(\mathbf{r} \mid \mathbf{a}=\tilde{\mathbf{a}}, \tilde{\theta}^{(i)}) \Pr[\mathbf{a}=\tilde{\mathbf{a}}]$$

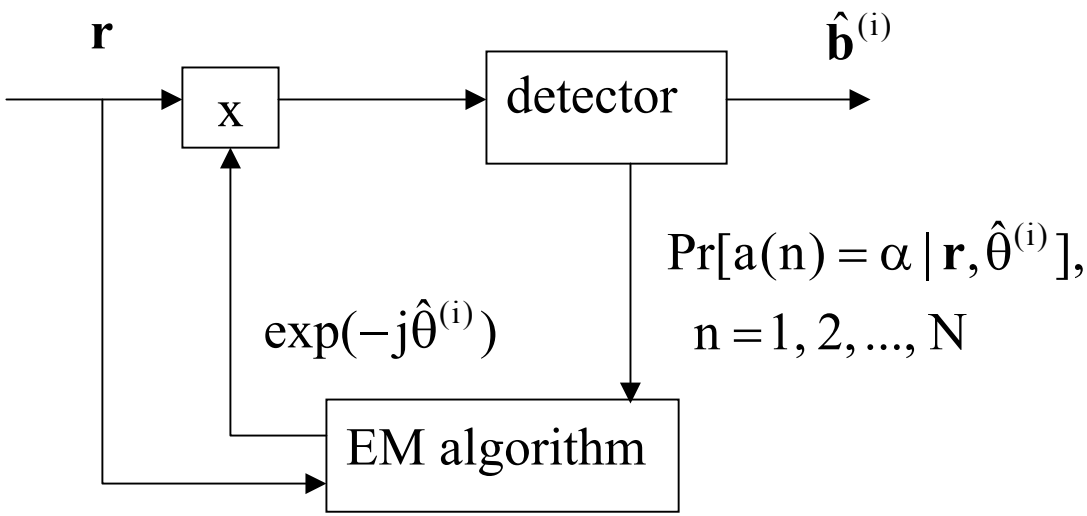
- som van ongeveer 2^k termen
- efficiënte (benaderde) berekening door detector, d.m.v. factorgraaf

in geval van niet-gecodeerde transmissie :

$$\Pr[a(n)=\alpha \mid \mathbf{r}, \tilde{\theta}^{(i)}] \propto p(r(n)|a(n)=\alpha, \tilde{\theta}^{(i)}) \Pr[a(n)=\alpha]$$

$$\propto \Pr[a(n)=\alpha] \exp\left(\frac{2}{N_0} \mathrm{Re}[\alpha^* r(n) \exp(-j\tilde{\theta}^{(i)}) \right)$$

EM algoritme

Voorbeeld (vervolg)

Ontvangerstructuur