Likelihood Functions

The likelihood function answers the question:

What does the sensor tell about the state $\mathbf{x}$ of the object?

(input: sensor data, sensor model)

- **ideal conditions, one object:** $P_D = 1$, $\rho_F = 0$

  at each time one measurement: $$p(\mathbf{z}_k|\mathbf{x}_k) = \mathcal{N}(\mathbf{z}_k; \mathbf{H}\mathbf{x}_k, \mathbf{R})$$

- **real conditions, one object:** $P_D < 1$, $\rho_F > 0$

  at each time n_k measurements $Z_k = \{\mathbf{z}_k^1, \ldots, \mathbf{z}_k^{n_k}\}$!

$$p(Z_k, n_k|\mathbf{x}_k) \propto (1 - P_D)\rho_F + P_D \sum_{j=1}^{n_k} \mathcal{N}\big(\mathbf{z}_k^j;\, \mathbf{H}\mathbf{x}_k,\, \mathbf{R}\big)$$

Bayes Filtering for: $P_D < 1, \rho_F > 0$, well-separated objects

state $\mathbf{x}_k$, ***current* data** $Z_k = \{\mathbf{z}_k^j\}_{j=1}^{m_k}$, ***accumulated* data** $\mathcal{Z}^k = \{Z_k, \mathcal{Z}^{k-1}\}$

interpretation hypotheses E_k **for** Z_k

object not detected, $1 - P_D$
$\mathbf{z}_k \in Z_k$ from object, P_D
$\Big\}$ $m_k + 1$ interpretations

interpretation histories H_k **for** $\mathcal{Z}^k$

- tree structure: $H_k = (E_{H_k}, H_{k-1}) \in \mathcal{H}^k$
- current: E_{H_k}, *pre*histories: H_{k-i}

$$p(\mathbf{x}_k|\mathcal{Z}^k) = \sum_{H_k} p(\mathbf{x}_k, H_k|\mathcal{Z}^k) = \sum_{H_k} \underbrace{p(H_k|\mathcal{Z}^k)}_{\text{weight!}} \underbrace{p(\mathbf{x}_k|H_k, \mathcal{Z}^k)}_{\textit{given } H_k\text{: unique}}$$

'mixture' density

Closer look: $P_D < 1, \rho_F > 0$, well-separated targets

filtering (at time t_{k-1}): $p(\mathbf{x}_{k-1}|\mathcal{Z}^{k-1}) = \sum_{H_{k-1}} p_{H_{k-1}} \mathcal{N}\big(\mathbf{x}_{k-1};\, \mathbf{x}_{H_{k-1}}, \mathbf{P}_{H_{k-1}}\big)$

prediction (for time t_k):

$$p(\mathbf{x}_k|\mathcal{Z}^{k-1}) = \int d\mathbf{x}_{k-1}\, p(\mathbf{x}_k|\mathbf{x}_{k-1})\, p(\mathbf{x}_{k-1}|\mathcal{Z}^{k-1}) \qquad \text{(MARKOV model)}$$

$$= \sum_{H_{k-1}} p_{H_{k-1}} \mathcal{N}\big(\mathbf{x}_k;\, \mathbf{F}\mathbf{x}_{H_{k-1}}, \mathbf{F}\mathbf{P}_{H_{k-1}}\mathbf{F}^\top + \mathbf{D}\big) \qquad \text{(IMM also possible)}$$

measurement likelihood:

$$p(Z_k, m_k|\mathbf{x}_k) = \sum_{j=0}^{m_k} p(\mathcal{Z}_k|E_k^j, \mathbf{x}_k, m_k)\, P(E_k^j|x_k, m_k) \qquad (E_k^j\text{: interpretations})$$

$$\propto (1 - P_D)\,\rho_F + P_D \sum_{j=1}^{m_k} \mathcal{N}\big(\mathbf{z}_k^j;\, \mathbf{H}\mathbf{x}_k,\, \mathbf{R}\big) \qquad (\mathbf{H}, \mathbf{R}, P_D, \rho_F)$$

filtering (at time t_k):

$$p(\mathbf{x}_k|\mathcal{Z}^k) \propto p(Z_k, m_k|\mathbf{x}_k)\, p(\mathbf{x}_k|\mathcal{Z}^{k-1}) \qquad \text{(BAYES' rule)}$$

$$= \sum_{H_k} p_{H_k} \mathcal{N}\big(\mathbf{x}_k;\, \mathbf{x}_{H_k}, \mathbf{P}_{H_k}\big) \qquad \text{(Exploit product formula)}$$

Problem: Growing Memory Disaster:

m data, N hypotheses $\to$ N^{m+1} continuations

radical solution: mono-hypothesis approximation

- **gating:** Exclude competing data with $||\boldsymbol{\nu}^i_{k|k-1}|| > \lambda$!

  $\to$ Kalman filter (KF)

  + very simple, – λ too small: loss of target measurement

- Force a **unique interpretation** in case of a conflict!

  look for *smallest statistical distance:* $\min_i ||\boldsymbol{\nu}^i_{k|k-1}||$

  $\to$ Nearest-Neighbor filter (NN)

  + one hypothesis, – hard decision, – not adaptive

- **global combining:** Merge all hypotheses!

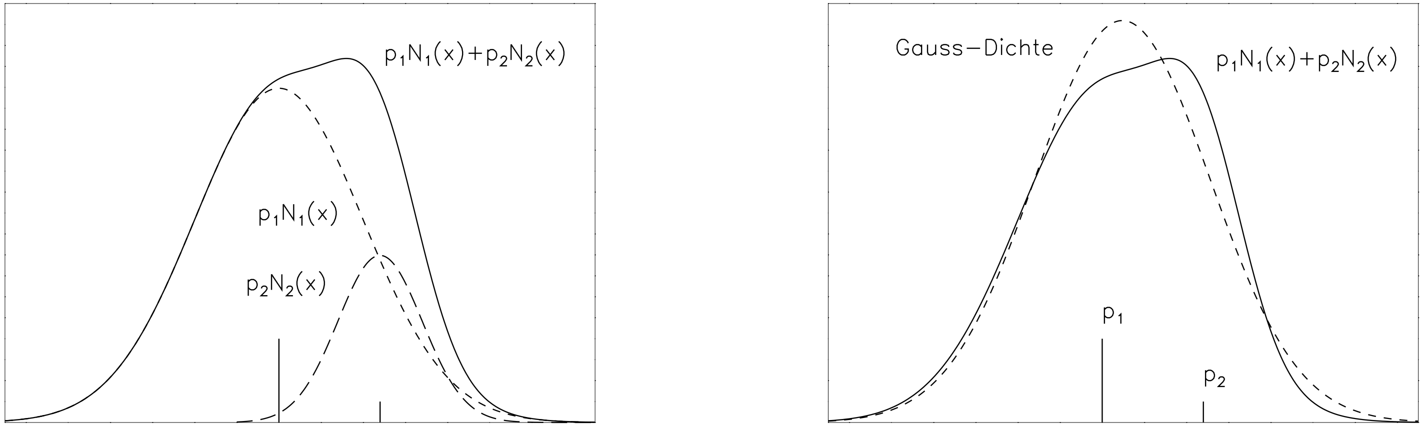
  $\to$ PDAF, JPDAF filter

  + all data, + adaptive, – reduced applicability

Moment Matching: Approximate an arbitrary pdf

$p(x)$ with $\mathbb{E}[x] = \mathbf{x}$, $\mathbb{C}[x] = \mathbf{P}$ by $p(x) \approx \mathcal{N}\big(x;\, \mathbf{x},\, \mathbf{P}\big)$!

here especially: $p(x) = \sum_H p_H \mathcal{N}(x; \mathbf{x}_H, \mathbf{P}_H)$ (normal mixtures)



$$\mathbf{x} = \sum_H p_H \, \mathbf{x}_H$$

$$\mathbf{P} = \sum_H p_H \left\{ \mathbf{P}_H + \overbrace{(\mathbf{x}_H - \mathbf{x})(\mathbf{x}_H - \mathbf{x})^\top}^{\text{spread term}} \right\}$$

PDAF Filter: formally analogous to Kalman Filter

Filtering (scan $k-1$): $p(\mathbf{x}_{k-1}|\mathcal{Z}^{k-1}) = \mathcal{N}(\mathbf{x}_{k-1};\, \mathbf{x}_{k-1|k-1}, \mathbf{P}_{k-1|k-1})$ (→ initiation)

prediction (scan k): $p(\mathbf{x}_k|\mathcal{Z}^{k-1}) \approx \mathcal{N}(\mathbf{x}_k;\, \mathbf{x}_{k|k-1}, \mathbf{P}_{k|k-1})$ (like Kalman)

Filtering (scan k): $p(\mathbf{x}_k|\mathcal{Z}^{k}) \approx \sum_{j=0}^{m_k} p_k^j \mathcal{N}(\mathbf{x}_k;\, \mathbf{x}_{k|k}^j, \mathbf{P}_{k|k}^j)$ $\approx \mathcal{N}(\mathbf{x}_k;\, \mathbf{x}_{k|k}, \mathbf{P}_{k|k})$

$$\mathbf{x}_{k|k}^j = \begin{cases} \mathbf{x}_{k|k-1} & j=0 \\ \mathbf{x}_{k|k-1} + \mathbf{W}_k \boldsymbol{\nu}_k^j & j\neq 0 \end{cases} \qquad \mathbf{P}_{k|k}^j = \begin{cases} \mathbf{P}_{k|k-1} & j=0 \\ \mathbf{P}_{k|k-1} - \mathbf{W}_k \mathbf{S}_k \mathbf{W}_k^\top & j\neq 0 \end{cases}$$

$$\boldsymbol{\nu}_k^j = \underbrace{\mathbf{z}_k^j - \mathbf{H}\mathbf{x}_k}_{\text{innovation}}, \qquad \mathbf{W}_k = \underbrace{\mathbf{P}_{k|k-1}\mathbf{H}^\top \mathbf{S}_k^{-1}}_{\text{gain matrix}}, \qquad \mathbf{S}_k = \underbrace{\mathbf{H}\mathbf{P}_{k|k-1}\mathbf{H}^\top + \mathbf{R}_k}_{\text{innovation covariance}}$$

$$p_k^j = \underbrace{\frac{p_k^{j*}}{\sum_j p_k^{j*}}}_{\text{Gewichte}}, \qquad p_k^{j*} = \begin{cases} (1-P_D)\,\rho_F & j=0 \\ \frac{P_D}{\sqrt{|2\pi \mathbf{S}_{H_k}|}}\, \mathrm{e}^{-\frac{1}{2}\boldsymbol{\nu}_{H_k}^\top \mathbf{S}_{H_k} \boldsymbol{\nu}_{H_k}} & j\neq 0 \end{cases}$$

PDAF Filter: formally analog to Kalman Filter

Filtering (scan $k-1$): $p(\mathbf{x}_{k-1}|\mathcal{Z}^{k-1}) = \mathcal{N}(\mathbf{x}_{k-1};\, \mathbf{x}_{k-1|k-1}, \mathbf{P}_{k-1|k-1})$ ($\rightarrow$ initiation)

prediction (scan k): $p(\mathbf{x}_k|\mathcal{Z}^{k-1}) \approx \mathcal{N}(\mathbf{x}_k;\, \mathbf{x}_{k|k-1}, \mathbf{P}_{k|k-1})$ (like Kalman)

Filtering (scan k): $p(\mathbf{x}_k|\mathcal{Z}^{k}) \approx \sum_{j=0}^{m_k} p_k^j\, \mathcal{N}(\mathbf{x}_k;\, \mathbf{x}_{k|k}^j, \mathbf{P}_{k|k}^j) \approx \mathcal{N}(\mathbf{x}_k;\, \mathbf{x}_{k|k}, \mathbf{P}_{k|k})$

$$\boldsymbol{\nu}_k = \boxed{\textstyle\sum_{j=0}^{m_k} p_k^j\, \boldsymbol{\nu}_k^j}, \qquad \boldsymbol{\nu}_k^j = \mathbf{z}_k^j - \mathbf{H}\mathbf{x}_{k|k-1}$$

combined innovation

$$\mathbf{W}_k = \mathbf{P}_{k|k-1}\mathbf{H}^\top\mathbf{S}_k^{-1}, \qquad \mathbf{S}_k = \mathbf{H}\mathbf{P}_{k|k-1}\mathbf{H}^\top + \mathbf{R}_k$$

Kalman gain matrix

$$p_k^j = p_k^{i*} \,/\, \textstyle\sum_j p_k^{j*}, \qquad p_k^{j*} = \begin{cases} (1-P_D)\,\rho_F \\ \frac{P_D}{\sqrt{|2\pi \mathbf{S}_{H_k}|}}\, \mathrm{e}^{-\frac{1}{2}\boldsymbol{\nu}_{H_k}^\top \mathbf{S}_{H_k} \boldsymbol{\nu}_{H_k}} \end{cases}$$

weighting factors

$$\mathbf{x}_k = \mathbf{x}_{k|k-1} + \mathbf{W}_k\, \boldsymbol{\nu}_k$$

(Filtering Update: Kalman)

$$\mathbf{P}_k = \mathbf{P}_{k|k-1} - (1-p_k^0)\, \mathbf{W}_k \mathbf{S} \mathbf{W}_k^\top$$

(Kalman part)

$$+\, \mathbf{W}_k \Big\{ \textstyle\sum_{j=0}^{m_k} p_k^j\, \boldsymbol{\nu}_k^j \boldsymbol{\nu}_k^{j\top} - \boldsymbol{\nu}_k \boldsymbol{\nu}_k{}^\top \Big\} \mathbf{W}_k^\top$$

(Spread of Innovations)

The qualitative shape of $p(\mathbf{x}_k|\mathcal{Z}^k)$ is often much simpler than its correct representation: *a few pronounced modes*

adaptive solution: nearly optimal approximation

The qualitative shape of $p(\mathbf{x}_k|\mathcal{Z}^k)$ is often much simpler than its correct representation: *a few pronounced modes*

adaptive solution: nearly optimal approximation

- **individual gating:** Exclude *irrelevant data*!

  before continuing existing track hypotheses H_{k-1}

  $\rightarrow$ *limiting case:* KALMAN filter (KF)

The qualitative shape of $p(\mathbf{x}_k|\mathcal{Z}^k)$ is often much simpler than its correct representation: *a few pronounced modes*

adaptive solution: nearly optimal approximation

- **individual gating:** Exclude *irrelevant data*!

  before continuing existing track hypotheses H_{k-1}

  $\rightarrow$ *limiting case:* KALMAN filter (KF)

- **pruning:** Kill hypotheses of very *small weight*!

  after calculating the weights p_{H_k}, before filtering

  $\rightarrow$ *limiting case:* Nearest Neighbor filter (NN)

The qualitative shape of $p(\mathbf{x}_k|\mathcal{Z}^k)$ is often much simpler than its correct representation: *a few pronounced modes*

adaptive solution: nearly optimal approximation

- **individual gating:** Exclude *irrelevant data*!

  before continuing existing track hypotheses H_{k-1}

  $\rightarrow$ *limiting case:* KALMAN filter (KF)

- **pruning:** Kill hypotheses of very *small weight*!

  after calculating the weights p_{H_k}, before filtering

  $\rightarrow$ *limiting case:* Nearest Neighbor filter (NN)

- **local combining:** Merge *similar hypotheses*!

  after the complete calculation of the pdfs

  $\rightarrow$ *limiting case:* PDAF (global combining)

Successive Local Combining

Partial sums of *similar* densities $\rightarrow$ moment matching:

$$\sum_{H_k \in \mathcal{H}^{k*}} p_{H_k} \mathcal{N}(\mathbf{x}_k; \mathbf{x}_{H_k}, \mathbf{P}_{H_k}) \approx p_{H_k^*} \mathcal{N}(\mathbf{x}_k; \mathbf{x}_{H_k^*}, \mathbf{P}_{H_k^*})$$

$\mathcal{H}^{k*} \subset \mathcal{H}^k \quad \rightarrow \quad H_k^*$: *effective* hypothesis

Successive Local Combining

Partial sums of *similar* densities $\to$ moment matching:

$$\sum_{H_k \in \mathcal{H}^{k*}} p_{H_k} \mathcal{N}(\mathbf{x}_k; \mathbf{x}_{H_k}, \mathbf{P}_{H_k}) \approx p_{H_k^*} \mathcal{N}(\mathbf{x}_k; \mathbf{x}_{H_k^*}, \mathbf{P}_{H_k^*})$$

$\mathcal{H}^{k*} \subset \mathcal{H}^k \quad \to \quad H_k^*$: *effective* hypothesis

similarity: $\quad d(H_1, H_2) < \mu \quad$ mit (z.B.):

$$d(H_1, H_2) = (\mathbf{x}_{H_1} - \mathbf{x}_{H_2})^\top (\mathbf{P}_{H_1} + \mathbf{P}_{H_2})^{-1} (\mathbf{x}_{H_1} - \mathbf{x}_{H_2})$$

Start: Hypothesis of highest weight $H_1 \to$ search similar hypothesis ($p_H \searrow$) $\to$ merge: $(H_1, H) \succ H_1^* \to$ continue search ($p_H \searrow$) ...

$\to$ **restart:** hypothesis with next to highest weight $H_2 \to \ldots$

Successive Local Combining

Partial sums of *similar* densities $\rightarrow$ moment matching:

$$\sum_{H_k \in \mathcal{H}^{k*}} p_{H_k} \mathcal{N}(\mathbf{x}_k; \mathbf{x}_{H_k}, \mathbf{P}_{H_k}) \approx p_{H_k^*} \mathcal{N}(\mathbf{x}_k; \mathbf{x}_{H_k^*}, \mathbf{P}_{H_k^*})$$

$\mathcal{H}^{k*} \subset \mathcal{H}^k \quad \rightarrow \quad H_k^*$: *effective* hypothesis

similarity: $\quad d(H_1, H_2) < \mu \quad$ mit (z.B.):

$$d(H_1, H_2) = (\mathbf{x}_{H_1} - \mathbf{x}_{H_2})^\top (\mathbf{P}_{H_1} + \mathbf{P}_{H_2})^{-1} (\mathbf{x}_{H_1} - \mathbf{x}_{H_2})$$

Start: Hypothesis of highest weight $H_1 \rightarrow$ search similar hypothesis ($p_H \searrow$) $\rightarrow$ merge: $(H_1, H) \succ H_1^* \rightarrow$ continue search ($p_H \searrow$) $\ldots$
$\rightarrow$ **restart:** hypothesis with next to highest weight $H_2 \rightarrow \ldots$

- In many cases: good approximations $\rightarrow$ *quasi-optimality*
- PDAF, JPDAF: $\mathcal{H}^{k*} = \mathcal{H}^k \rightarrow$ *limited applicability*
- robustness $\rightarrow$ *detail mostly irrelevant*

Retrodiction for Gaussian Mixtures

wanted: $p(\mathbf{x}_l|\mathcal{Z}^k) \longleftarrow p(\mathbf{x}_{l+1}|\mathcal{Z}^k)$ for $l < k$

$$p(\mathbf{x}_l|\mathcal{Z}^k) = \sum_{H_k} p(\mathbf{x}_l, H_k|\mathcal{Z}_k) = \sum_{H_k} \underbrace{p(\mathbf{x}_l|H_k, \mathcal{Z}^k)}_{\text{no ambiguities!}} \underbrace{p(H_k|\mathcal{Z}^k)}_{\text{filtering!}}$$

Calculation of $p(\mathbf{x}_l|H_k, \mathcal{Z}^k)$ as in case of $P_D = 1$, $\rho_F = 0$!

$$p(\mathbf{x}_l|H_k, \mathcal{Z}^k) = \mathcal{N}\Big(\mathbf{x}_l;\, \mathbf{x}_{H_k}(l|k),\, \mathbf{P}_{H_k}(l|k)\Big)$$

with parameters given by Rauch-Tung-Striebel formulae:

$$\mathbf{x}_{H_k}(l|k) = \mathbf{x}_{H_k}(l|l) + \mathbf{W}_{H_k}(l|k)\,(\mathbf{x}_{H_k}(l+1|k) - \mathbf{x}_{H_k}(l+1|l))$$
$$\mathbf{P}_{H_k}(l|k) = \mathbf{P}_{H_k}(l|l) + \mathbf{W}_{H_k}(l|k)\,(\mathbf{P}_{H_k}(l+1|k) - \mathbf{P}_{H_k}(l+1|l))\,\mathbf{W}_{H_k}(l|k)^\top$$

gain matrix: $\mathbf{W}_{H_k}(l|k) = \mathbf{P}_{H_k}(l|l)\mathbf{F}_{l+1|l}^\top \mathbf{P}_{H_k}(l+1|l)^{-1}$

Retrodiction of Hypotheses' Weights

Consider approximation: neglect RTS step!

$$p(\mathbf{x}_l|H_k, \mathcal{Z}^k) = \mathcal{N}\big(\mathbf{x}_l;\, \mathbf{x}_{H_k}(l|k),\, \mathbf{P}_{H_k}(l|k)\big) \approx \mathcal{N}\big(\mathbf{x}_l;\, \mathbf{x}_{H_k}(l|l),\, \mathbf{P}_{H_k}(l|l)\big)$$

$$p(\mathbf{x}_l|H_k, \mathcal{Z}^k) \approx \sum_{H_l} p^*_{H_l}\, \mathcal{N}\big(\mathbf{x}_l;\, \mathbf{x}_{H_k}(l|l),\, \mathbf{P}_{H_k}(l|l)\big)$$

with recursively defined weights: $p^*_{H_k} = p_{H_k}, \quad p^*_{H_l} = \sum p^*_{H_{l+1}}$

summation over all histories H_{l+1} with equal pre-histories!

- Strong sons strengthen weak fathers.
- Weak sons weaken even strong fathers.
- If all sons die, also the father must die.

Recapitulation: Detection Process for Sensors

Detector: receives signals and decides on object existence

Processor: processes detected signals and produces measurements

‘D’: detector detects an object

D: object actually existent

error of 1. kind: $P_{\mathrm{I}} = P(\neg\text{‘}D\text{’}|D)$

error of 2. kind: $P_{\mathrm{II}} = P(\text{‘}D\text{’}|\neg D)$

measure of detection performance: $P_D = P(\text{‘}D\text{’}|D)$

detector properties characterized by two parameters:

– detection probability $P_D = 1 - P_{\mathrm{I}}$
– false alarm probability $P_F = P_{\mathrm{II}}$

example (Swerling I model): $P_D = P_D(P_F, \text{SNR}) = P_F^{1/(1+\text{SNR})}$

detector design: Maximize detection probability P_D for a given, predefined false alarm probability P_F.

Track Extraction: Initiation of the PDF Iteration

***extraction of target tracks*: detection on a higher level of abstraction**

start: data sets $Z_k = \{\mathbf{z}_k^j\}_{j=1}^{m_k}$ (sensor performance: P_D, ρ_F, $\mathbf{R}$)

goal: Detect a target trajectory in a time series: $\mathcal{Z}^k = \{Z_i\}_{i=1}^k$!

at first simplifying assumptions:

- The targets in the sensors' field of view (FoV) are well-separated.
- The sensor data in the FoV in scan i are produced simultaneously.

Track Extraction: Initiation of the PDF Iteration

***extraction of target tracks*: detection on a higher level of abstraction**

start: data sets $Z_k = \{\mathbf{z}_k^j\}_{j=1}^{m_k}$ (sensor performance: P_D, ρ_F, $\mathbf{R}$)

goal: Detect a target trajectory in a time series: $\mathcal{Z}^k = \{Z_i\}_{i=1}^k$!

at first simplifying assumptions:

- The targets in the sensors' field of view (FoV) are well-separated.
- The sensor data in the FoV in scan i are produced simultaneously.

decision between two competing hypotheses:

h_1: Besides false returns $\mathcal{Z}^k$ contains also target measurements.

h_0: There is no target existing in the FoV; all data in $\mathcal{Z}^k$ are false.

statistical decision errors:

$P_1 = \text{Prob}(\textit{accept } h_1 | h_1)$ analogous to the sensors' P_D

$P_0 = \text{Prob}(\textit{accept } h_1 | h_0)$ analogous to the sensors' P_F

Practical Approach: Sequential Likelihood Ratio Test

Goal: Decide as fast as possible for given decision errors P_0, P_1!

Consider the ratio of the conditional probabilities $p(h_1|\mathcal{Z}^k)$, $p(h_0|\mathcal{Z}^k)$ and the likelihood ratio $\mathrm{LR}(k) = p(\mathcal{Z}^k|h_1)/p(\mathcal{Z}^k|h_0)$ as an intuitive decision function:

$$\frac{p(h_1|\mathcal{Z}^k)}{p(h_0|\mathcal{Z}^k)} = \frac{p(\mathcal{Z}^k|h_1)}{p(\mathcal{Z}^k|h_0)} \frac{p(h_1)}{p(h_0)} \qquad \text{a priori: } p(h_1) = p(h_0)$$

Practical Approach: Sequential Likelihood Ratio Test

Goal: Decide as fast as possible for given decision errors P_0, P_1!

Consider the ratio of the conditional probabilities $p(h_1|\mathcal{Z}^k)$, $p(h_0|\mathcal{Z}^k)$ and the likelihood ratio $\text{LR}(k) = p(\mathcal{Z}^k|h_1)/p(\mathcal{Z}^k|h_0)$ as an intuitive decision function:

$$\frac{p(h_1|\mathcal{Z}^k)}{p(h_0|\mathcal{Z}^k)} = \frac{p(\mathcal{Z}^k|h_1)}{p(\mathcal{Z}^k|h_0)} \frac{p(h_1)}{p(h_0)}$$

a priori: $p(h_1) = p(h_0)$

Starting from a time window with length $k = 1$, calculate the test function $\text{LR}(k)$ successively and compare it with *two* thresholds A, B:

If	$\text{LR}(k) < A$,	accept hypothesis h_0 (i.e. no target is existing)!
If	$\text{LR}(k) > B$,	accept hypothesis h_1 (i.e. target exists in FoV)!
If	$A < \text{LR}(k) < B$,	wait for new data Z_{k+1}, repeat with $\text{LR}(k+1)$!

Fraunhofer FKIE

Iterative Calculation of the Likelihood Ratio

$$\text{LR}(k) = \frac{p(\mathcal{Z}^k|h_1)}{p(\mathcal{Z}^k|h_0)} = \frac{\int d\mathbf{x}_k\, p(Z_k, m_k, \mathbf{x}_k, \mathcal{Z}^{k-1}|h_1)}{p(Z_k, m_k, \mathcal{Z}^{k-1}, h_0)}$$

Iterative Calculation of the Likelihood Ratio

$$\text{LR}(k) = \frac{p(\mathcal{Z}^k|h_1)}{p(\mathcal{Z}^k|h_0)} = \frac{\int d\mathbf{x}_k \, p(Z_k, m_k, \mathbf{x}_k, \mathcal{Z}^{k-1}|h_1)}{p(Z_k, m_k, \mathcal{Z}^{k-1}, h_0)} = \frac{\int d\mathbf{x}_k \, p(Z_k, m_k|\mathbf{x}_k) \, p(\mathbf{x}_k|\mathcal{Z}^{k-1}, h_1) \, p(\mathcal{Z}^{k-1}|h_1)}{|\text{FoV}|^{-m_k} \, p_F(m_k) \, p(\mathcal{Z}^{k-1}|h_0)}$$

Iterative Calculation of the Likelihood Ratio

$$\mathsf{LR}(k) = \frac{p(\mathcal{Z}^k|h_1)}{p(\mathcal{Z}^k|h_0)} = \frac{\int d\mathbf{x}_k\, p(Z_k, m_k, \mathbf{x}_k, \mathcal{Z}^{k-1}|h_1)}{p(Z_k, m_k, \mathcal{Z}^{k-1}, h_0)} = \frac{\int d\mathbf{x}_k\, p(Z_k, m_k|\mathbf{x}_k)\, p(\mathbf{x}_k|\mathcal{Z}^{k-1}, h_1)\, p(\mathcal{Z}^{k-1}|h_1)}{|\mathsf{FoV}|^{-m_k}\, p_F(m_k)\, p(\mathcal{Z}^{k-1}|h_0)}$$

$$= \frac{\int d\mathbf{x}_k\, p(Z_k, m_k|\mathbf{x}_k, h_1)\, p(\mathbf{x}_k|\mathcal{Z}^{k-1}, h_1)}{|\mathsf{FoV}|^{-m_k}\, p_F(m_k)}\, \mathsf{LR}(k-1)$$

basic idea: iterative calculation!

Iterative Calculation of the Likelihood Ratio

$$\mathsf{LR}(k) = \frac{p(\mathcal{Z}^k|h_1)}{p(\mathcal{Z}^k|h_0)} = \frac{\int d\mathbf{x}_k\, p(Z_k, m_k, \mathbf{x}_k, \mathcal{Z}^{k-1}|h_1)}{p(Z_k, m_k, \mathcal{Z}^{k-1}, h_0)} = \frac{\int d\mathbf{x}_k\, p(Z_k, m_k|\mathbf{x}_k)\, p(\mathbf{x}_k|\mathcal{Z}^{k-1}, h_1)\, p(\mathcal{Z}^{k-1}|h_1)}{|\mathsf{FoV}|^{-m_k}\, p_F(m_k)\, p(\mathcal{Z}^{k-1}|h_0)}$$

$$= \frac{\int d\mathbf{x}_k\, p(Z_k, m_k|\mathbf{x}_k, h_1)\, p(\mathbf{x}_k|\mathcal{Z}^{k-1}, h_1)}{|\mathsf{FoV}|^{-m_k}\, p_F(m_k)}\, \mathsf{LR}(k-1)$$

basic idea: iterative calculation!

Let $H_k = \{E_k, H_{k-1}\}$ be an interpretation history of the time series $\mathcal{Z}^k = \{Z_k, \mathcal{Z}^{k-1}\}$.

$E_k = E_k^0$: target was not detected, $E_k = E_k^j$: $\mathbf{z}_k^j \in Z_k$ is a target measurement.

$$p(\mathbf{x}_k|\mathcal{Z}^{k-1}, h_1) = \sum_{H_{k-1}} p(\mathbf{x}_k|H_{k-1}\mathcal{Z}^{k-1}, h_1)\, p(H_{k-1}|\mathcal{Z}^{k-1}, h_1)$$ The standard MHT prediction!

$$p(Z_k, m_k|\mathbf{x}_k, h_1, h_1) = \sum_{E_k} p(Z_k, E_k|\mathbf{x}_k, h_1)$$ The standard MHT likelihood function!

The calculation of the likelihood ratio is just *a by-product of Bayesian MHT tracking*.

Fraunhofer FKIE

Iteration Formula for $\mathsf{LR}(k) = p(\mathcal{Z}^k|h_1)/p(\mathcal{Z}^k|h_0)$

initiation: $k = 0, \quad \mathbf{j}_0 = 0, \quad \lambda_{\mathbf{j}_0} = 1$

recursion: $$\mathsf{LR}(k+1) = \sum_{\mathbf{j}_{k+1}} \lambda_{\mathbf{j}_{k+1}} = \sum_{j_{k+1}=0}^{m_{k+1}} \sum_{\mathbf{j}_k} \lambda_{j_{k+1}\mathbf{j}_k} \lambda_{\mathbf{j}_k}$$

with: $$\lambda_{j_{k+1}\mathbf{j}_k} = \begin{cases} 1 - P_D & \text{for} \quad j_{k+1} = 0 \\ \frac{P_D}{\rho_F} \mathcal{N}(\boldsymbol{\nu}_{j_{k+1}\mathbf{j}_k}, \mathbf{S}_{j_{k+1}\mathbf{j}_k}) & \text{for} \quad j_{k+1} \neq 0 \end{cases}$$

convenient notation: with $\mathbf{j}_k = (j_k, \ldots, j_1)$ let $\sum_{\mathbf{j}_k} \lambda_{\mathbf{j}_k} = \sum_{j_k=0}^{m_k} \cdots \sum_{j_1=0}^{m_1} \lambda_{j_k \ldots j_1}$

Iteration Formula for $\mathsf{LR}(k) = p(\mathcal{Z}^k|h_1)/p(\mathcal{Z}^k|h_0)$

initiation: $k = 0, \quad \mathbf{j}_0 = 0, \quad \lambda_{\mathbf{j}_0} = 1$

recursion: $$\mathsf{LR}(k+1) = \sum_{\mathbf{j}_{k+1}} \lambda_{\mathbf{j}_{k+1}} = \sum_{j_{k+1}=0}^{m_{k+1}} \sum_{\mathbf{j}_k} \lambda_{j_{k+1}\mathbf{j}_k} \lambda_{\mathbf{j}_k}$$

with: $$\lambda_{j_{k+1}\mathbf{j}_k} = \begin{cases} 1 - P_D & \text{for} \quad j_{k+1} = 0 \\ \frac{P_D}{\rho_F} \mathcal{N}(\boldsymbol{\nu}_{j_{k+1}\mathbf{j}_k}, \mathbf{S}_{j_{k+1}\mathbf{j}_k}) & \text{for} \quad j_{k+1} \neq 0 \end{cases}$$

innovation: $\boldsymbol{\nu}_{j_{k+1}\mathbf{j}_k} = \mathbf{z}_{j_{k+1}} - \mathbf{H}_{j_{k+1}} \mathbf{x}_{\mathbf{j}_{k+1|k}}$

innov. cov.: $\mathbf{S}_{j_{k+1}\mathbf{j}_k} = \mathbf{H}_{j_{k+1}} \mathbf{P}_{\mathbf{j}_{k+1|k}} \mathbf{H}_{j_{k+1}}^\top + \mathbf{R}_{j_{k+1}}$

state update: $\mathbf{x}_{\mathbf{j}_{k+1|k}} = \mathbf{F}_{j_{k+1}} \mathbf{x}_{\mathbf{j}_k} \qquad \mathbf{x}_{\mathbf{j}_k} = \mathbf{x}_{\mathbf{j}_{k|k-1}} + \mathbf{W}_{j_k \mathbf{j}_{k-1}} \boldsymbol{\nu}_{j_k, \mathbf{j}_{k-1}}$

covariances: $\mathbf{P}_{\mathbf{j}_{k+1|k}} = \mathbf{F}_{j_{k+1}} \mathbf{P}_{\mathbf{j}_k} \mathbf{F}_{j_{k+1}}^\top + \mathbf{D}_{j_{k+1}} \qquad \mathbf{P}_{\mathbf{j}_k} = \mathbf{P}_{\mathbf{j}_{k|k-1}} - \mathbf{W}_{j_k \mathbf{j}_{k-1}} \mathbf{S}_{j_k \mathbf{j}_{k-1}} \mathbf{W}_{j_k \mathbf{j}_{k-1}}^\top$

Iteration Formula for $\mathsf{LR}(k) = p(\mathcal{Z}^k|h_1)/p(\mathcal{Z}^k|h_0)$

initiation: $k = 0, \quad \mathbf{j}_0 = 0, \quad \lambda_{\mathbf{j}_0} = 1$

recursion: $$\mathsf{LR}(k+1) = \sum_{\mathbf{j}_{k+1}} \lambda_{\mathbf{j}_{k+1}} = \sum_{j_{k+1}=0}^{m_{k+1}} \sum_{\mathbf{j}_k} \lambda_{j_{k+1}\mathbf{j}_k} \lambda_{\mathbf{j}_k}$$

with: $$\lambda_{j_{k+1}\mathbf{j}_k} = \begin{cases} 1 - P_D & \text{for} \quad j_{k+1} = 0 \\ \frac{P_D}{\rho_F} \mathcal{N}(\boldsymbol{\nu}_{j_{k+1}\mathbf{j}_k}, \mathbf{S}_{j_{k+1}\mathbf{j}_k}) & \text{for} \quad j_{k+1} \neq 0 \end{cases}$$

innovation: $\boldsymbol{\nu}_{j_{k+1}\mathbf{j}_k} = \mathbf{z}_{j_{k+1}} - \mathbf{H}_{j_{k+1}} \mathbf{x}_{\mathbf{j}_{k+1|k}}$

innov. cov.: $\mathbf{S}_{j_{k+1}\mathbf{j}_k} = \mathbf{H}_{j_{k+1}} \mathbf{P}_{\mathbf{j}_{k+1|k}} \mathbf{H}_{j_{k+1}}^\top + \mathbf{R}_{j_{k+1}}$

state update: $\mathbf{x}_{\mathbf{j}_{k+1|k}} = \mathbf{F}_{j_{k+1}} \mathbf{x}_{\mathbf{j}_k}$ $\qquad \mathbf{x}_{\mathbf{j}_k} = \mathbf{x}_{\mathbf{j}_{k|k-1}} + \mathbf{W}_{j_k\mathbf{j}_{k-1}} \boldsymbol{\nu}_{j_k,\mathbf{j}_{k-1}}$

covariances: $\mathbf{P}_{\mathbf{j}_{k+1|k}} = \mathbf{F}_{j_{k+1}} \mathbf{P}_{\mathbf{j}_k} \mathbf{F}_{j_{k+1}}^\top + \mathbf{D}_{j_{k+1}}$ $\qquad \mathbf{P}_{\mathbf{j}_k} = \mathbf{P}_{\mathbf{j}_{k|k-1}} - \mathbf{W}_{j_k\mathbf{j}_{k-1}} \mathbf{S}_{j_k\mathbf{j}_{k-1}} \mathbf{W}_{j_k\mathbf{j}_{k-1}}^\top$

Exercise 7.1 Show that this recursion formulae for calculating the decision function is true.

Sequential Track Extraction: Discussion

- $\mathsf{LR}(k)$ is given by a growing number of summands, each related to a particular interpretation history. The tuple $\{\lambda_{\mathbf{j}_k}, \mathbf{x}_{\mathbf{j}_k} \mathbf{P}_{\mathbf{j}_k}\}$ is called a *sub-track*.

Sequential Track Extraction: Discussion

- $\mathsf{LR}(k)$ is given by a growing number of summands, each related to a particular interpretation history. The tuple $\{\lambda_{\mathbf{j}_k}, \mathbf{x}_{\mathbf{j}_k} \mathbf{P}_{\mathbf{j}_k}\}$ is called a *sub-track*.

- For mitigating growing memory problems all approximations discussed for track maintenance can be used if they do not significantly affect $\mathsf{LR}(k)$:

  - *individual gating:* Exclude data not likely to be associated.
  - *pruning:* Kill sub-tacks contributing marginally to the test function.
  - *local combining:* Merge similar sub tracks:

$$\{\lambda_i, \mathbf{x}_i, \mathbf{P}_i\}_i \to \{\lambda, \mathbf{x}, \mathbf{P}\} \qquad \text{with:} \quad \lambda = \textstyle\sum_i \lambda_i,$$
$$\mathbf{x} = \tfrac{1}{\lambda} \textstyle\sum_i \lambda_i \mathbf{x}_i, \quad \mathbf{P} = \tfrac{1}{\lambda} \textstyle\sum_i \lambda_i [\mathbf{P}_i + (\mathbf{x}_i - \mathbf{x})(\ldots)^\top].$$

Sequential Track Extraction: Discussion

- $\mathsf{LR}(k)$ is given by a growing number of summands, each related to a particular interpretation history. The tuple $\{\lambda_{\mathbf{j}_k}, \mathbf{x}_{\mathbf{j}_k} \mathbf{P}_{\mathbf{j}_k}\}$ is called a *sub-track*.
- For mitigating growing memory problems all approximations discussed for track maintenance can be used if they do not significantly affect $\mathsf{LR}(k)$:
  - *individual gating:* Exclude data not likely to be associated.
  - *pruning:* Kill sub-tacks contributing marginally to the test function.
  - *local combining:* Merge similar sub tracks:

$$\{\lambda_i, \mathbf{x}_i, \mathbf{P}_i\}_i \to \{\lambda, \mathbf{x}, \mathbf{P}\} \qquad \text{with:} \quad \lambda = \textstyle\sum_i \lambda_i,$$
$$\mathbf{x} = \tfrac{1}{\lambda} \textstyle\sum_i \lambda_i \mathbf{x}_i, \quad \mathbf{P} = \tfrac{1}{\lambda} \textstyle\sum_i \lambda_i [\mathbf{P}_i + (\mathbf{x}_i - \mathbf{x})(\ldots)^\top].$$

- The LR test ends with a decision in favor of or against the hypotheses: h_0 (no target) or h_1 (target existing). Intuitive interpretation of the thresholds!

track extraction at t_k: Decide in favor of h_1!

initiation of pdf iteration (track maintenance):

Normalize coefficients $\lambda_{\mathbf{j}_k}$: $p_{\mathbf{j}_k} = \frac{\lambda_{\mathbf{j}_k}}{\sum_{\mathbf{j}_k} \lambda_{\mathbf{j}_k}}$!

$$(\lambda_{\mathbf{j}_k}, \mathbf{x}_{\mathbf{j}_k}, \mathbf{P}_{\mathbf{j}_k}) \quad \rightarrow \quad p(\mathbf{x}_k|\mathcal{Z}^k) = \sum_{\mathbf{j}_k} p_{\mathbf{j}_k} \, \mathcal{N}\big(\mathbf{x}_k;\, \mathbf{x}_{\mathbf{j}_k},\, \mathbf{P}_{\mathbf{j}_k}\big)$$

Continue track extraction with the remaining sensor data!

sequential LR test for track monitoring:

After deciding in favor of h_1 reset $\mathrm{LR}(0) = 1$! Calculate $\mathrm{LR}(k)$ from $p(\mathbf{x}_k|\mathcal{Z}^k)$!

track confirmation: $\mathrm{LR}(k) > \frac{P_1}{P_0}$: reset $\mathrm{LR}(0) = 1$!

track deletion: $\mathrm{LR}(k) < \frac{1-P_1}{1-P_0}$; ev. track re-initiation

DEMONSTRATION (simulated)

DEMONSTRATION (simulated)

Exercise 7.2 (voluntary)

Simulate a detection process with a given P_D, target measurements with a given $\mathbf{R}$, a detection process with a given P_D and realize the track extraction procedure.

Generalization to Target Cluster (Perfect Resolution)

Scheme directly extendable to clusters consisting of n targets, *if n is known*!

principal approach in case of *unknown* n:

1. Start with sensor measurements Z_1.
2. Assume for a target cluster $n \leq N$! A-prioriliy: $P(n) = \frac{1}{N}$
3. hypothesis h_n: there exist n targets; the data set Z_1 contains at least one target measurement; h_0: no target existing at all
4. Consider the following ratio (at least 1, at most N targets):

$$\frac{p(h_1 \vee \ldots \vee h_N|\mathcal{Z}^k)}{p(h_0|\mathcal{Z}^k)} = \frac{\sum_{n=1}^N p(h_n|\mathcal{Z}^k)}{p(h_0|\mathcal{Z}^k)} = \sum_{n=1}^N \frac{p(\mathcal{Z}^k|h_n)}{p(\mathcal{Z}^k|h_0)} \frac{p(h_n)}{p(h_0)}$$

Generalization to Target Cluster (Perfect Resolution)

Scheme directly extendable to clusters consisting of n targets, *if* n *is known*!

principal approach in case of *unknown* n:

1. Start with sensor measurements Z_1.

2. Assume for a target cluster $n \leq N$! A-priorily: $P(n) = \frac{1}{N}$

3. hypothesis h_n: there exist n targets; the data set Z_1 contains at least one target measurement; h_0: no target existing at all

4. generalized LR test function: $\mathrm{LR}(k) = \frac{1}{N} \sum_{n=1}^{N} \frac{p(\mathcal{Z}^k|h_n)}{p(\mathcal{Z}^k|h_0)}$

5. Calculate $\mathrm{LR}_n(k) = p(\mathcal{Z}^k|h_n)/p(\mathcal{Z}^k|h_0)$ in analogy to $n = 1$.

6. 'Cardinality' of having n objects in the cluster: $c_k(n) = \frac{\mathrm{LR}_n(k)}{\sum_{n=1}^{N} \mathrm{LR}_n(k)}$

DEMONSTRATION (simulated)

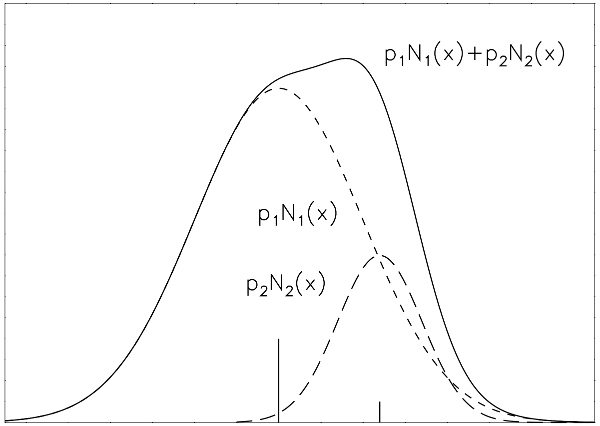
ABRAHAM WALD (1902-1950) Austro-Hungarian mathematician who contributed to decision theory, geometry, and econometrics; founded the theory of economic equilibria in Oskar Morgenstern's institut in Vienna: "Berechnung der Ausschaltung von Saisonschwankungen" (Springer Verlag, 1936) the basis of Game Theory: Morgenstern, John von Neumann, John Forbes Nash (1994: Nobel price with Reinhard Selten, Bonn University) → sensor management! Founder of statistical sequential analysis in WW II. 1950 plenary talk at the International Congress of Mathematicians ICM, Cambridge (Mass.): "Basic ideas of a general theory of statistical decision rules" (1900: Hilbert's 23 Problems).

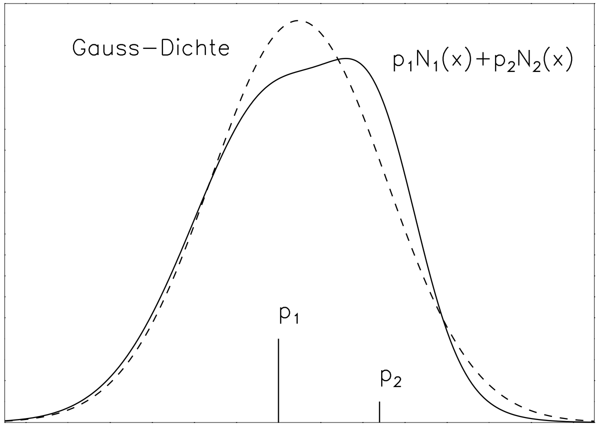
Student and friend: Jacob Wolfowitz (statistician, information theory), classical text book: "Coding Theorems of Information Theory" (1978). Posthumous attack by Ronald Fisher: "an incompetent book on statistics", passionately defended by Jerzy Neyman as imminent a statistician as Fisher.

Moment Matching: Approximate an arbitrary pdf

$p(x)$ with $\mathbb{E}[x] = \mathbf{x}$, $\mathbb{C}[x] = \mathbf{P}$ by $p(x) \approx \mathcal{N}\big(x;\, \mathbf{x},\, \mathbf{P}\big)$!

here especially: $p(x) = \sum_i p_i \mathcal{N}(x; \mathbf{x}_i, \mathbf{P}_i)$ (GAUSSian mixtures)





Exercise Show:

$$\mathbf{x} = \sum_i p_i \, \mathbf{x}_i$$

$$\mathbf{P} = \sum_i p_i \left\{ \mathbf{P}_i + \overbrace{(\mathbf{x}_i - \mathbf{x})(\mathbf{x}_i - \mathbf{x})^\top}^{\text{spread term}} \right\}$$

Fraunhofer FKIE

2nd Order Approximation:

here: $p(x) = \sum_i p_H \mathcal{N}(x; \mathbf{x}_i, \mathbf{P}_i) \approx \mathcal{N}(x; \mathbb{E}_p[x], \mathbb{C}_p[x])$

$$\mathbb{E}_p[x] = \int dx\, x\, p(x) = \sum_i p_i \int dx\, x\, \mathcal{N}(x; \mathbf{x}_i, \mathbf{P}_i) = \sum_i p_i\, \mathbf{x}_i =: \mathbf{x}$$

$$\mathbb{C}_p[x] = \int dx\, p(x)\, (x - \mathbb{E}_p[x])(x - \mathbb{E}_p[x])^\top = \sum_i p_i \int dx\, (x - \mathbf{x})(x - \mathbf{x})^\top \mathcal{N}(x; \mathbf{x}_i, \mathbf{P}_i)$$

$$= \sum_i p_i \int dx\, \{(x - \mathbf{x})(x - \mathbf{x})^\top - 2(x - \mathbf{x}_i)(\mathbf{x}_i - \mathbf{x})^\top\} \mathcal{N}(x; \mathbf{x}_i, \mathbf{P}_i)$$

since we have: $\int dx\, (x - \mathbf{x}_i)(\mathbf{x}_i - \mathbf{x})^\top \mathcal{N}(x; \mathbf{x}_i, \mathbf{P}_i) = 0$

$$= \sum_i p_i \int dx\, \{x x^\top - 2x\mathbf{x}_i^\top + \mathbf{x}_i\mathbf{x}_i^\top + \mathbf{x}_i\mathbf{x}_i^\top - 2\mathbf{x}_i\mathbf{x}^\top + \mathbf{x}\mathbf{x}^\top\} \mathcal{N}(x; \mathbf{x}_i, \mathbf{P}_i)$$

$$= \sum_i p_i \int dx\, \{(x - \mathbf{x}_i)(x - \mathbf{x}_i)^\top + (\mathbf{x}_i - \mathbf{x})(\mathbf{x}_i - \mathbf{x})^\top\} \mathcal{N}(x; \mathbf{x}_i, \mathbf{P}_i)$$

$$= \sum_i p_i \{\mathbf{P}_i + (\mathbf{x}_i - \mathbf{x})(\mathbf{x}_i - \mathbf{x})^\top\} = \mathbf{P}$$