INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

# Odhad stavu stochastických systémů

Miroslav Šimandl (FAV, ZČU v Plzni)

8. prosince 2010

Tato prezentace je spolufinancována Evropským sociálním fondem a státním rozpočtem České republiky.

# Problém odhadu
## Základní pojmy

**Odhadování:** proces hledání hodnoty veličiny z nepřímých, nepřesných a nejistých pozorování.

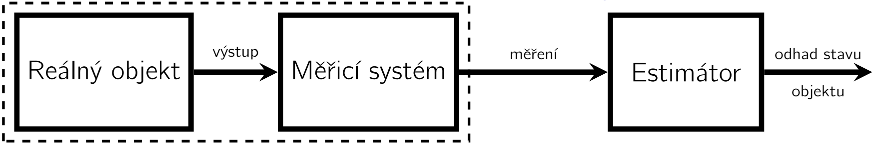
**Odhad:** nalezená hodnota veličiny
V případě, že hledaná veličina je náhodná proměnná, je za odhad často považováno nalezení hustoty pravděpodobnosti této veličiny nebo bodový odhad.

**Estimátor:** systém generující odhad.

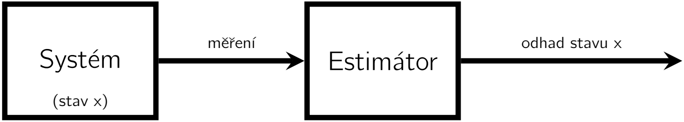
**Rozhodování:** výběr jedné možnosti z množiny alternativ.

# Problém odhadu
## Základní pojmy

odhadování neměřitelného stavu reálného objektu




blokové schéma popisující úlohu odhadu



# Problém odhadu
## Základní pojmy

**Filtrace:** odhad veličin nebo stochastických procesů ze zašuměných dat

**Sledování:** cílem je sledovat atributy objektu a měřicí zařízení je mimo objekt

**Navigace:** cílem je sledovat atributy objektu a měřicí zařízení je součástí objektu

# Aplikace odhadování

- sledování
- navigace
- řídicí systémy
- řízení strojů a technologických procesů
- zpracování signálu
- zpracování obrazu
- komunikace (zpracování řeči, komunikace člověk-stroj)

# Problém odhadu
## Aplikace odhadování

- biomedicínské inženýrství
- operační výzkum
- geofyzikální problémy
- ekonometrické systémy
- demografické systémy
- dopravní systémy
- předpověď počasí

**Estimace prováděná Gaussem** (1777-1855) - proces vyvozování informace o parametrech, např. pohyb nebeských těles, přesně predikovatelný pohyb

**Estimace ve smyslu Kalmana** (1960 - Kalmanův filtr) - vyvozování informace o stavu objektu, který se pohybuje ne zcela přesně predikovatelným způsobem.

# Gaussovy postřehy

- Jestliže je pozorování absolutně přesné, stačí $n$ pozorování pro $n$ parametrů. Následující pozorování by mělo potvrdit, ale ne korigovat odhad.
- Pozorování je pouze aproximací skutečnosti, mělo by se využít více pozorování než je minimální počet s ohledem na počet neznámých parametrů. Odhady by se měly korigovat následnými pozorováními a je vhodné použít všechny pozorování k získání odhadu.

# Základní principy odhadu parametrů

- Existuje základní popis modelu systému s neznámými parametry, které mají být odhadnuty.
- Redundantní data jsou požadována pro redukci efektu chyb měření.
- Je nutné zajistit, aby rozdíl mezi pozorovanými hodnotami a predikovanými z odhadů byl co nejmenší.
- Kombinace počáteční znalosti a následného pozorování vede ke koncepci rekurzivního výpočtu.
- Nepřesnost pozorování vyžaduje statistické přístupy.

# Nejmenší čtverce (Gauss-Legendre)

- Odhad **konstantních parametrů** modelu minimalizací kvadrátu rozdílu výstupu modelu a měřeného výstupu systému.
- Pro gaussovské šumy je miminalizace součtu kvadrátu reziduí ekvivalentní maximalizaci věrohodnostní funkce parametrů.
- Později použito pro odhad náhodné veličiny.

Wiener-Hopfův problém, Kolmogorovův problém (30. až 40. léta 20. století)

- **Odhadované veličiny nejsou konstantní**, ale časově proměnné.
- Cíl je odhadnout náhodný stacionární proces.
- Odhad procesu je obdržen ve frekvenční oblasti.

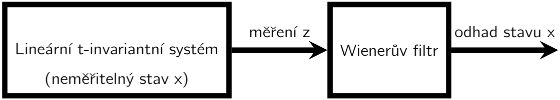
# Kalman-Bucy problém (od roku 1960)

- Odhaduje se **náhodný proces**, ne nutně stacionární.
- **Popis** - lineární diferenciální nebo diferenční rovnice se známými parametry.
- **Řešení** v časové oblasti odpovídá minimalizaci ve střední kvadratické chybě.
- **Estimátor** je nazýván Kalman-Bucy filtr nebo Kalmanův filtr.

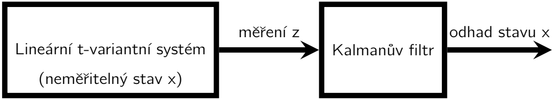
# Problém nelineárního odhadu (od roku 1970)

- Odhaduje se náhodný proces generovaný
  - nelineárním systémem
  - lineárním systémem s negaussovskými poruchami
- Výsledkem je nalezení
  - podmíněné hustoty pravděpodobnosti stavu
  - podmíněný bodový odhad stavu

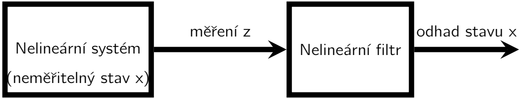
### Odhad stacionárního náh. procesu $x(t)$ pomocí Wienerova filtru




### Odhad náhodného procesu $x(t)$ pomocí Kalmanova filtru




### Odhad náhodného procesu $x(t)$ pomocí nelineárního filtru



Uvažujme jako přirozený základ pro diskusi o odhadu parametrů a odhadu stavu následující rovnici

$$z_k = y_k + v_k \quad k = 0, 1, \ldots$$

kde $z_k$ je měření na systému dostupné v čase $t_k$
$y_k$ je výstup systému v čase $t_k$
$v_k$ reprezentuje šum měření neboli poruchu v čase $t_k$.

Problém je najít odhad $y_k$ z měření $z_k$ pro všechna $k$.

Primitivní odhad by mohl být

$$\hat{y}_k \triangleq z_k .$$

Chyba odhadu by v tomto případě byla

$$\tilde{y}_k \triangleq y_k - \hat{y}_k = z_k - v_k - z_k = -v_k .$$

Cíl estimační teorie je navrhnout takový postup při stanovení odhadu, který bude redukovat chybu odhadu na menší, než kterou dává primitivní estimátor

Problém: chyba $\tilde{y}_k$ nemůže být určena explicitně.
Proto je třeba znát více o signálu $y_k$ - předpokládejme nejdříve, že

$$y_k = \theta$$

$$z_k = \theta + v_k \quad k = 0, 1, 2, ..., N.$$

Odhad $\theta$ můžeme provést

$$(\sum_{k=0}^{N} z_k)/(N+1) = \theta + (\sum_{k=0}^{N} v_k)/(N+1).$$

Rozumný odhad $\theta$

$$\hat{\theta} \triangleq \left(\sum_{k=0}^{N} z_k\right)/(N+1)$$

a chyba odhadu

$$\tilde{\theta} = \theta - \hat{\theta} = -\left(\sum_{k=0}^{N} v_k\right)/(N+1).$$

Jestliže je však $v_k$ konstanta, pak k žádnému zlepšení odhadu nedojde, přestože jsme věděli více o výstupu $y_k$ a použili statistický přístup.

Rozšiřme úvahy na situaci popsanou rovnicí

nelineární problém estimace parametrů $\Theta$

$$y_k = h_k(\Theta) \quad k = 0, 1, 2, \ldots$$

kde $h_k(\cdot)$ je známá vektorová funkce

nebo

lineární problém estimace parametrů $\Theta$

$$y_k = H_k.\Theta \quad k = 0, 1, 2, \ldots$$

kde $H_k$ je známá matice příslušných dimenzí.

# Nelineární estimace stavu

Mnohdy však není možné uvažovat neznámé jako konstantní veličiny a pak

$$y_k = h_k(x_k).$$

Vývoj vektorové proměnné $x_k$ je popsán diferenční rovnicí

$$x_{k+1} = f_k(x_k) + w_k$$

kde $f_k(\cdot)$ je známá vektorová funkce.
$w_k$ je neznámý vektor reprezentující šum v čase $t_k$.

# Lineární estimace stavu nebo parametrů

speciální případ

$$x_{k+1} = F_k x_k + w_k$$

$$y_k = H_k x_k$$

kde $F_k$ a $H_k$ jsou známé matice příslušných dimenzí.
Pro

$$x_{k+1} = x_k = \Theta$$

nebo

$$x_{k+1} = f_k(x_k)$$

tedy stavový šum je nulový pro všechna $k$ dostáváme problém estimace parametrů, $x_0 = \Theta$.

## závěr

- strukturální vlastnosti
- jednoduchost při aplikaci
- problémy s kvalitativním ohodnocením přesnosti odhadu

K popisu neurčitosti se tradičně používá počtu pravděpodobnosti.

Aby měření poskytovalo nejvíce informací, měl by šum $v^N = (v_0, v_1, ..., v_N)$ splňovat

$$p(v^N) = p(v_0, v_1, ..., v_N) = p(v_0).p(v_1)...p(v_N)$$

Tedy, aby stochastický proces $\{v_k\}$ byl bílý šum.

# Stavový šum

- Stavový šum $w_k$ nesmí vykazovat žádnou závislost do minulosti.

$$p(w^k) = p(w_0, w_1, \ldots, w_k) = p(w_0).p(w_1)\ldots p(w_k)$$

- Nezávislost stavového šumu, šumu měření a počátečního stavu $x_0$,
- Poznamenejme, že v tomto případě $x_k$ pro $k = 0, 1, 2, \ldots$ reprezentuje markovský proces.

# Problém estimace parametrů

Mějme vektor měření $z_k$ popsaný rovnicí

$$z_k = h_k(\Theta) + v_k \quad k = 0, 1, 2, \ldots, N,$$

kde $E[v_k] = 0, E[v_k v_l^T] = R_k \delta_{kj}, \quad \delta_{kj} = 1, k = j; \delta_{kj} = 0, \quad k \neq j$
$h_k(\cdot)$ je známá vektorová funkce.

Cílem je určit odhad neznámých konstantních parametrů $\Theta$.

# Problém estimace stavu

Mějme vektor měření $z_k$ popsaný

$$z_k = h_k(x_k) + v_k \quad k = 0, 1, 2, ..., N$$

kde $E[v_k] = 0$, $E[v_k v_j^T] = R_k \delta_{kj}$, $h_k(\cdot)$ je známá nelineání funkce a vektor stavu se vyvíjí podle rovnice

$$x_{k+1} = f_k(x_k) + w_k \quad k = 0, 1, 2, ..., N$$

kde $E[w_k] = 0, E[w_k w_j^T] = Q_k \delta_{kj}$, $f_k(\cdot)$ je známá nelineání funkce.

Cílem je odhadnout stav $x_k$.

- Uvažujme náhodné vektory $\mathbf{x}$ a $\mathbf{z}$,
- vektor $\mathbf{x}$ je neznámý,
- vektor $\mathbf{z}$ reprezentuje data, vztažená k $\mathbf{x}$,
- předpokládejme apriorní informaci o neznámé $\mathbf{x}$.

- Cílem je najít pro dané $\mathbf{z}$ odhad $\hat{\mathbf{x}}(\mathbf{z})$ náhodné veličiny $\mathbf{x}$ tak, aby chyba odhadu $\tilde{\mathbf{x}} = \mathbf{x} - \hat{\mathbf{x}}(\mathbf{z})$ byla malá.

Definujme střední kvadratickou chybu jako střední hodnotu kvadrátu eukleidovské normy chyby odhadu $\tilde{\mathbf{x}}$

$$J = \mathrm{E}\left[\tilde{\mathbf{x}}^{\mathrm{T}}\tilde{\mathbf{x}}\right] = \mathrm{E}\left[(\mathbf{x} - \hat{\mathbf{x}}(\mathbf{z}))^{\mathrm{T}}(\mathbf{x} - \hat{\mathbf{x}}(\mathbf{z}))\right].$$

Odhad podle střední kvadratické chyby je speciální případ tzv. Bayesova odhadu, kde se hledá odhad minimalizující Bayesovo riziko

$$J = \mathrm{E}\left[C(\tilde{\mathbf{x}})\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} C(\mathbf{x} - \hat{\mathbf{x}}(\mathbf{z}))p_{\mathrm{xz}}(\mathbf{x}, \mathbf{z})\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{z}$$

kde $C()$ je penalizační, kriteriální funkce.

Interpretace: použití veškeré dostupné informace k nalezení optimálního odhadu $\hat{\mathbf{x}}$ minimalizující kritérium střední kvadratické chyby.

dostupná informace:

1. Existuje pouze informace o statistice $\mathbf{x}$, o pravděpodobnostním rozložení.
2. Existuje apriorní informace o statistikách $\mathbf{x}$ a $\mathbf{z}$, a navíc je k dispozici hodnota měření $\mathbf{z}$.

Předpokládejme, že nemáme žádné měření **z**, ale známe apriorní statistiky neznámé **x**, tedy že známe $p_{\mathrm{x}}(\mathbf{x})$.

$$J(\hat{\mathbf{x}}) = \mathrm{E}\left[(\mathbf{x}-\hat{\mathbf{x}})^{\mathrm{T}}(\mathbf{x}-\hat{\mathbf{x}})\right] = \mathrm{E}\left[\mathbf{x}^{\mathrm{T}}\mathbf{x}\right] - 2\hat{\mathbf{x}}^{\mathrm{T}}\mathrm{E}\left[\mathbf{x}\right] + \hat{\mathbf{x}}^{\mathrm{T}}\hat{\mathbf{x}}.$$

Derivací předchozího výrazu je zřejmé

$$\frac{\partial J}{\partial \hat{\mathbf{x}}} = -2\mathrm{E}\left[\mathbf{x}\right] + 2\hat{\mathbf{x}},$$

a tedy

$$\hat{\mathbf{x}}_{MS} = \mathrm{E}\left[\mathbf{x}\right].$$

Nejlepší konstanta jako odhad $\hat{\mathbf{x}}_{MS}$ je $\mathrm{E}\left[\mathbf{x}\right]$.

# Klasická teorie odhadu
## MS odhad při znalosti pouze apriorní informace

Odhad je nestranný

$$\mathrm{E}\left[\tilde{\mathbf{x}}\right] = \mathrm{E}\left[\mathbf{x} - \hat{\mathbf{x}}\right] = \mathrm{E}\left[\mathbf{x}\right] - \hat{\mathbf{x}} = \mathrm{E}\left[\mathbf{x}\right] - \mathrm{E}\left[\mathbf{x}\right] = 0.$$

$$\mathrm{E}\left[\tilde{\mathbf{x}}\right] = 0$$

nebo ekvivalentně

$$\mathrm{E}\left[\hat{\mathbf{x}}\right] = \mathrm{E}\left[\mathbf{x}\right].$$

Míru důvěry v odhad vyjadřuje kovariance chyby

$$\mathbf{P}_{\tilde{x}} = \mathrm{E}\left[(\tilde{\mathbf{x}} - \mathrm{E}[\tilde{\mathbf{x}}])(\tilde{\mathbf{x}} - \mathrm{E}[\tilde{\mathbf{x}}])^{\mathrm{T}}\right] = \mathrm{E}\left[\tilde{\mathbf{x}}\tilde{\mathbf{x}}^{\mathrm{T}}\right] = \mathrm{E}\left[(\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})^{\mathrm{T}}\right]$$
$$= \mathrm{E}\left[(\mathbf{x} - \mathrm{E}[\mathbf{x}])(\mathbf{x} - \mathrm{E}[\mathbf{x}])^{\mathrm{T}}\right]$$

což je apriorní kovarianční matice veličiny $\mathbf{x}$.

Apriorní kovarianční matice $\mathbf{P}_{x}$ je tedy rovna kovarianci chyby

$$\mathbf{P}_{\tilde{x}} = \mathbf{P}_{x}.$$

# Klasická teorie odhadu

## MS odhad využitím apriorní informace i měření

Předpokládejme, že navíc ke statistice $\mathbf{x}$ známe i statistiku $\mathbf{z}$, která je v nějakém vtahu k $\mathbf{x}$. Tento vztah obecně vyjádříme pomocí sdružené hustoty pravděpodobnosti.

$$p_{\mathrm{xz}}(\mathbf{x}, \mathbf{z}).$$

Optimální odhad podle střední kvadratické chyby závisí na (známých) měřeních, tedy $\hat{\mathbf{x}}(\mathbf{z})$. Tento odhad můžeme nazývat aposteriorním odhadem $\mathbf{x}$.
Abychom nalezli odhad $\hat{\mathbf{x}}(\mathbf{z})$, uvažujme kritérium

$$J = \tilde{\mathbf{x}}^{\mathrm{T}}\tilde{\mathbf{x}} = \iint_{-\infty}^{\infty} (\mathbf{x} - \hat{\mathbf{x}}(\mathbf{z}))^{\mathrm{T}} (\mathbf{x} - \hat{\mathbf{x}}(\mathbf{z})) p_{\mathrm{xz}}(\mathbf{x}, \mathbf{z}) \mathrm{d}\mathbf{x} \mathrm{d}\mathbf{z}.$$

Použitím Bayesova pravidla

$$p_{\mathrm{x},\mathrm{z}}(\mathbf{x}, \mathbf{z}) = p_{\mathrm{x}|\mathrm{z}}(\mathbf{x}|\mathbf{z}) p_{\mathrm{z}}(\mathbf{z})$$

můžeme předchozí vztah přepsat následujícím způsobem

$$J = \int_{-\infty}^{\infty} p_{\mathrm{z}}(\mathbf{z}) \int_{-\infty}^{\infty} \left(\mathbf{x} - \hat{\mathbf{x}}(\mathbf{z})\right)^{\mathrm{T}} \left(\mathbf{x} - \hat{\mathbf{x}}(\mathbf{z})\right) p_{\mathrm{x}|\mathrm{z}}(\mathbf{x}|\mathbf{z}) \mathrm{d}\mathbf{x} \mathrm{d}\mathbf{z}.$$

V tomto případě lze minimalizovat pouze vnitřní integrál

$$J_V = \int_{-\infty}^{\infty} \left(\mathbf{x} - \hat{\mathbf{x}}(\mathbf{z})\right)^{\mathrm{T}} \left(\mathbf{x} - \hat{\mathbf{x}}(\mathbf{z})\right) p_{\mathrm{x}|\mathrm{z}}(\mathbf{x}|\mathbf{z}) \mathrm{d}\mathbf{x}.$$

Optimální odhad lze najít takto

$$
\begin{aligned}
J_V &= \mathrm{E}\left[(\mathbf{x} - \hat{\mathbf{x}}(\mathbf{z}))^{\mathrm{T}} (\mathbf{x} - \hat{\mathbf{x}}(\mathbf{z})) \,\middle|\, \mathbf{z}\right] \\
&= \mathrm{E}\left[\mathbf{x}^{\mathrm{T}}, \mathbf{x}|\mathbf{z}\right] - 2\hat{\mathbf{x}}^{\mathrm{T}}(\mathbf{z})\mathrm{E}\left[\mathbf{x}|\mathbf{z}\right] + \hat{\mathbf{x}}^{\mathrm{T}}(\mathbf{z})\hat{\mathbf{x}}(\mathbf{z}).
\end{aligned}
$$

Derivací kritéria podle $\hat{\mathbf{x}}(\mathbf{z})$ dostaneme

$$
\frac{\partial J_V}{\partial \hat{\mathbf{x}}(\mathbf{z})} = -2\mathrm{E}\left[\mathbf{x}|\mathbf{z}]\right] + 2\hat{\mathbf{x}}(\mathbf{z}) = 0,
$$

tedy

$$
\hat{\mathbf{x}}_{MS}(\mathbf{z}) = \mathrm{E}\left[\mathbf{x}|\mathbf{z}\right].
$$

Pokud měření $z$ neobsahuje informaci o $\mathbf{x}$, tedy $p_{\mathrm{xz}}(\mathbf{x},\mathbf{z}) = p_{\mathrm{x}}(\mathbf{x})p_{\mathrm{z}}(\mathbf{z})$ pak

$$\mathrm{E}\left[\mathbf{x}|\mathbf{z}\right] = \int \mathbf{x} p_{\mathrm{x}|\mathrm{z}}(\mathbf{x}|\mathbf{z})\mathrm{d}\mathbf{x} = \int \mathbf{x}\frac{p_{\mathrm{x,z}}(\mathbf{x},\mathbf{z})}{p_{\mathrm{z}}(\mathbf{z})}\mathrm{d}\mathbf{x} = \int \mathbf{x}\frac{p_{\mathrm{x}}(\mathbf{x})p_{\mathrm{z}}(\mathbf{z})}{p_{\mathrm{z}}(\mathbf{z})} = \mathrm{E}\left[\mathbf{x}\right].$$

Odhad ve smyslu podmíněné střední hodnoty je nestranný

$$\begin{aligned} 0 &= \mathrm{E}\left[\tilde{\mathbf{x}}\right] = \mathrm{E}\left[\mathrm{E}\left[\tilde{\mathbf{x}}|\mathbf{z}\right]\right] = \mathrm{E}\left[\mathrm{E}\left[\mathbf{x}-\hat{\mathbf{x}}|\mathbf{z}\right]\right] = \mathrm{E}\left[\mathrm{E}\left[\mathbf{x}-\mathrm{E}\left[\mathbf{x}|\mathbf{z}\right]|\mathbf{z}\right]\right] \\ &= \mathrm{E}\left[\mathrm{E}\left[\mathbf{x}|\mathbf{z}\right]-\mathrm{E}\left[\mathrm{E}\left[\mathbf{x}|\mathbf{z}\right]|\mathbf{z}\right]\right] = \mathrm{E}\left[\mathrm{E}\left[\mathbf{x}|\mathbf{z}\right]-\mathrm{E}\left[\mathbf{x}|\mathbf{z}\right]\right] = 0. \end{aligned}$$

Míra důvěry v odhad je opět dána kovariancí chyby, tedy

$$\begin{aligned}
\mathbf{P}_{\tilde{\mathbf{x}}} &= \mathrm{E}\left[\left(\tilde{\mathbf{x}} - \mathrm{E}\left[\tilde{\mathbf{x}}\right]\right)\left(\tilde{\mathbf{x}} - \mathrm{E}\left[\tilde{\mathbf{x}}\right]\right)^{\mathrm{T}}\right] = \mathrm{E}\left[\tilde{\mathbf{x}}\tilde{\mathbf{x}}^{\mathrm{T}}\right] \\
&= \mathrm{E}\left[\left(\mathbf{x} - \hat{\mathbf{x}}\right)\left(\mathbf{x} - \hat{\mathbf{x}}\right)^{\mathrm{T}}\right] = \mathrm{E}\left[\mathrm{E}\left[\left(\mathbf{x} - \mathrm{E}\left[\mathbf{x}|\mathbf{z}\right]\right)\left(\mathbf{x} - \mathrm{E}\left[\mathbf{x}|\mathbf{z}\right]\right)^{\mathrm{T}} \middle| \mathbf{z}\right]\right] \\
&= \mathrm{E}\left[\mathbf{P}_{\mathrm{x}|\mathrm{z}}\right]
\end{aligned}$$

A tedy

$$\mathbf{P}_{\tilde{\mathbf{x}}} = \mathrm{E}\left[\mathbf{P}_{\mathrm{x}|\mathrm{z}}\right].$$

Omezíme se nyní na speciální případ, kdy odhady $\hat{\mathbf{x}}(\mathbf{z})$, které vykazují lineární závislost na měření $\mathbf{z}$

$$\hat{\mathbf{x}}(\mathbf{z}) = \mathbf{A}\mathbf{z} + \mathbf{b},$$

kde $\mathbf{A}$ je konstantní matice, $\mathbf{A} \in \mathbb{R}^{n\times p}$ a $\mathbf{b} \in \mathbb{R}^{n}$.

Cílem je najít matici $\mathbf{A}$ a vektor $\mathbf{b}$ pro kritérium ve smyslu střední kvadratické chyby

$$J = \mathrm{E}\left[(\mathbf{x} - \hat{\mathbf{x}}(\mathbf{z}))^{\mathrm{T}}(\mathbf{x} - \hat{\mathbf{x}}(\mathbf{z}))\right].$$

Dosazením struktury estimátoru do kritéria dostaneme

$$
\begin{aligned}
J =& \mathrm{E}\left[(\mathbf{x}-\hat{\mathbf{x}})^{\mathrm{T}}(\mathbf{x}-\hat{\mathbf{x}})\right] = \operatorname{tr}\mathrm{E}\left[(\mathbf{x}-\hat{\mathbf{x}})(\mathbf{x}-\hat{\mathbf{x}})^{\mathrm{T}}\right] \\
=& \operatorname{tr}\mathrm{E}\left[[(\mathbf{x}-\mathrm{E}[\mathbf{x}])-(\mathbf{Az}+\mathbf{b}-\mathrm{E}[\mathbf{x}])][(\mathbf{x}-\mathrm{E}[\mathbf{x}])-(\mathbf{Az}+\mathbf{b}-\mathrm{E}[\mathbf{x}])]^{\mathrm{T}}\right]
\end{aligned}
$$

$$
\begin{aligned}
=& \operatorname{tr}\left[\mathbf{P}_{\mathrm{x}}+\mathbf{A}\left(\mathbf{P}_{\mathrm{z}}+\mathrm{E}[\mathbf{z}]\,\mathrm{E}[\mathbf{z}]^{\mathrm{T}}\right)\mathbf{A}^{\mathrm{T}}+(\mathbf{b}-\mathrm{E}[\mathbf{x}])(\mathbf{b}-\mathrm{E}[\mathbf{x}])^{\mathrm{T}}\right. \\
&\left.+\,2\mathbf{Az}((\mathbf{b}-\mathrm{E}[\mathbf{x}])^{\mathrm{T}}-2\mathbf{AP}_{\mathrm{zx}}\right].
\end{aligned}
$$

Pro výpočet minima potřebujeme řešit

$$\frac{\partial J}{\partial \mathbf{b}} = 2(\mathbf{b} - \mathrm{E}[\mathbf{x}]) + 2\mathbf{A}\mathrm{E}[\mathbf{z}] = 0,$$

$$\frac{\partial J}{\partial \mathbf{A}} = 2\mathbf{A}(\mathbf{P}_z + \mathrm{E}[\mathbf{z}]\,\mathrm{E}[\mathbf{z}]^{\mathrm{T}}) - 2\mathbf{P}_{xz} + 2\,(\mathbf{b} - \mathrm{E}[\mathbf{x}])\,\mathrm{E}[\mathbf{z}]^{\mathrm{T}} = 0.$$

Z první rovnice dostaneme hodnotu vektoru **b**

$$\mathbf{b} = \mathrm{E}[\mathbf{x}] - \mathbf{A}\mathrm{E}[\mathbf{z}]$$

a po dosazení do druhé rovnice obdržíme

$$\mathbf{A}\mathbf{P}_z - \mathbf{P}_{xz} = 0$$

z čehož plyne

$$\mathbf{A} = \mathbf{P}_{xz}\mathbf{P}_z^{-1}.$$

Nejlepší lineární estimátor ve smyslu střední kvadratické chyby

$$\hat{\mathbf{x}}_{LMS} = \mathrm{E}\left[\mathbf{x}\right] + \mathbf{P}_{\mathrm{xz}}\mathbf{P}_{\mathrm{z}}^{-1}(\mathbf{z} - \mathrm{E}\left[\mathbf{z}\right])$$

- Pro
  $$\hat{\mathbf{x}}_{MS}(\mathbf{z}) = \mathrm{E}\left[\mathbf{x}|\mathbf{z}\right]$$
  potřebujeme $p_{\mathrm{x}|\mathrm{z}}(\mathbf{x}|\mathbf{z})$.
- Pro
  $$\hat{\mathbf{x}}_{LMS}(\mathbf{z}) = \mathrm{E}\left[\mathbf{x}\right] + \mathbf{P}_{\mathrm{xz}}\mathbf{P}_{\mathrm{z}}^{-1}(\mathbf{z} - \mathrm{E}\left[\mathbf{z}\right])$$
  potřebujeme momenty 1. a 2. řádu $\mathrm{E}\left[\mathbf{x}\right]$, $\mathrm{E}\left[\mathbf{z}\right]$ a $\mathbf{P}_{\mathrm{xz}}$, $\mathbf{P}_{\mathrm{z}}$.

Snadno lze dokázat, že $\hat{\mathbf{x}}_{LMS}$ je nestranný

$$\mathrm{E}\left[\tilde{\mathbf{x}}\right] = \mathrm{E}\left[\mathbf{x} - \hat{\mathbf{x}}_{LMS}\right] = \mathrm{E}\left[\mathbf{x}\right] - \mathrm{E}\left[\mathbf{x}\right] = 0$$

a kovariance chyby je

$$\mathbf{P}_{\tilde{x}} = \mathrm{E}\left[\left(\tilde{\mathbf{x}} - \mathrm{E}\left[\tilde{\mathbf{x}}\right]\right)\left(\tilde{\mathbf{x}} - \mathrm{E}\left[\tilde{\mathbf{x}}\right]\right)^{\mathrm{T}}\right] = \mathrm{E}\left[\tilde{\mathbf{x}}\tilde{\mathbf{x}}^{\mathrm{T}}\right] = \mathbf{P}_{x} - \mathbf{P}_{xz}\mathbf{P}_{z}^{-1}\mathbf{P}_{zx}.$$

# Klasická teorie odhadu
## Optimální odhad podle maximální věrohodnosti - ML

- Předpokládejme, že $\mathbf{x}$ je deterministická neznámá veličina a $\mathbf{z}$ je stochastická veličina. O veličině $\mathbf{x}$ není nic dalšího známo. Apriorní informace existuje o veličině $\mathbf{z}$ ve smyslu existence a znalosti parametrické hustoty $p_{\mathrm{z}}(\mathbf{z};\mathbf{x})$.
- Mějme měření $\mathbf{z}$, odhad ve smyslu maximální věrohodnosti(ML) $\hat{\mathbf{x}}^{ML}$ je hodnota $\mathbf{x}$, která maximalizuje $p_{\mathrm{z}}(\mathbf{z};\mathbf{x})$, tedy věrohodnost, že $\mathbf{x}$ je promítnuto v měření $\mathbf{z}$.
- $p_{\mathrm{z}}(\mathbf{z};\mathbf{x})$ nazývá věrohodnostní funkce. Za věrohodnostní funkci lze též považovat $\ln p_{\mathrm{z}}(\mathbf{z};\mathbf{x})$.

Odhad ve smyslu maximální věrohodnosti může být nalezen z

$$\frac{\partial p_{\mathrm{z}}(\mathbf{z};\mathbf{x})}{\partial \mathbf{x}}\Big|_{\mathrm{x}=\hat{\mathrm{x}}} = 0$$

nebo

$$\frac{\partial \ln p_{\mathrm{z}}(\mathbf{z};\mathbf{x})}{\partial \mathbf{x}}\Big|_{\mathrm{x}=\hat{\mathrm{x}}} = 0$$

což jsou věrohodnostní rovnice.

Předpokládejme, že existuje lineární vztah mezi měřením $\mathbf{z}$ a neznámou veličinou $\mathbf{x}$

$$\mathbf{z} = \mathbf{H}\mathbf{x} + \mathbf{v},$$

kde $\mathbf{v}$ je šum měření s normálním rozložením

$$p_{\mathrm{v}}(\mathbf{v}) = \frac{1}{\sqrt{(2\pi)^p|\mathrm{R}|}} e^{-\frac{1}{2}\mathrm{v}^{\mathrm{T}}\mathrm{R}^{-1}\mathrm{v}}.$$

Pak je věrohodnostní funkce určena vztahem

$$p_{\mathrm{z}}(\mathbf{z};\mathbf{x}) = p_{\mathrm{v}}(\mathbf{z} - \mathbf{H}\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^p|\mathrm{R}|}} e^{-\frac{1}{2}(\mathrm{z}-\mathrm{Hx})^{\mathrm{T}}\mathrm{R}^{-1}(\mathrm{z}-\mathrm{Hx})}.$$

Maximalizace $p_z(\mathbf{z};\mathbf{x})$ může být provedena maximalizací $\ln p_z(\mathbf{z};\mathbf{x})$ nebo minimalizací

$$J = \frac{1}{2}(\mathbf{z}-\mathbf{H}\mathbf{x})^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{z}-\mathbf{H}\mathbf{x})$$

. Derivací $J$ podle $\mathbf{x}$ dostaneme

$$\frac{\partial J}{\partial \mathbf{x}} = \mathbf{H}^{\mathrm{T}}R^{-1}(\mathbf{z}-\mathbf{H}\mathbf{x}) = 0.$$

takže optimální odhad je

$$\hat{\mathbf{x}}_{ML} = (\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{z}$$

# Klasická teorie odhadu

## ML odhad pro lineární měření a gaussovské poruchy

Odhad je nestranný

$$\begin{aligned}\tilde{\mathbf{x}} =& \mathbf{x} - \hat{\mathbf{x}} = \mathbf{x} - (\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{H}\mathbf{x} + \mathbf{v}) \\ =& -(\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{v},\end{aligned}$$

a tedy

$$\mathrm{E}\left[\tilde{\mathbf{x}}\right] = -(\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathrm{E}\left[\mathbf{v}\right] = 0.$$

Kovariance chyby odhadu je

$$\mathbf{P}_{\tilde{x}} = \mathrm{E}\left[\left(\tilde{\mathbf{x}} - \mathrm{E}\left[\tilde{\mathbf{x}}\right]\right)\left(\tilde{\mathbf{x}} - \mathrm{E}\left[\tilde{\mathbf{x}}\right]\right)^{\mathrm{T}}\right] = (\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})^{-1}.$$

Využitím $\mathbf{P}_{\tilde{\mathrm{x}}}$ můžeme optimální odhad zapsat

$$\hat{\mathbf{x}}^{ML} = \mathbf{P}_{\tilde{\mathrm{x}}}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{z}.$$

Absence apriorní informace o $\mathbf{x}$ může být vyjádřena vztahy

$$\bar{\mathbf{x}} = 0$$
$$\mathbf{P}_{\mathrm{x}} \to \infty.$$

Pak jsou vztahy pro střední kvadratickou chybu a maximální věrohodnost v limitě stejné.

# Bayesovský přístup

1. Uvažujme náhodný vektor $z = [z_1, z_2, ..., z_N]^T$ s hustotou pravděpodobnosti $p(z; \Theta)$, kde $\Theta = [\Theta_1, \Theta_2, ..., \Theta_l]^T$ je vektor parametrů.
2. Při klasickém přístupu k problému odhadu $\Theta$ považujeme parametry za neznámé konstanty a k závěrům o $\Theta$ použijeme pouze $z$ a tvar rozdělení $z$.
3. Při bayesovském přístupu k závěrům o parametru $\Theta$, tentokrát chápaného jako náhodná proměnná použijeme kromě toho ještě apriorní informaci o parametru $\Theta$, kterou máme k dispozici nezávisle na realizaci $z$.
4. Apriorní informace se vyjadřuje předpokladem, že $\Theta$ je náhodný vektor s jistým rozložením. Tato informace může mít objektivní i subjektivní charakter.

Aposteriorní hustota pravděpodobnosti $\Theta$ pro dané $z^k$ je podle Bayesova pravidla (Thomas Bayes 1761)

$$p(\Theta \mid z^k) = \frac{p(z^k \mid \Theta).p(\Theta)}{p(z^k)}$$

kde $p(z^k) = \int p(z^k \mid \Theta).p(\Theta) d\Theta$.
Za předpokladu, že $\Theta$ a $v^k = (v_0, v_1, \ldots, v_k)$ jsou nezávislé, $p(z^k \mid \Theta)$ je známa z $p(v^k)$, protože vycházíme z platnosti modelu měření

$$z_k = h_k(\Theta) + v_k \quad k = 0, 1, 2, \ldots$$

a tedy

$$p(z^k \mid \Theta) = \prod_{j=0}^{k} p_{v_j}(z_j - h(\Theta)).$$

Určení podmíněné hustoty pravděpodobnosti $p(x_k \mid z^l)$ nebo odhadu $x_k$ na základě pozorování $z^l$

$k = l$: filtrace

$l > k$: vyhlazování

$k > l$: predikce

$$x_{k+1} = f_k(x_k) + w_k \quad k = 0, 1, 2, \ldots$$

Náhodný proces $\{w_k\}$ je bílý šum se známou hustotou pravděpodobnosti $p(w_k)$ a $p(x_0)$ je známa.

$$z_k = h_k(x_k) + v_k \quad k = 0, 1, 2, \ldots$$

Náhodný proces $\{v_k\}$ je bílý šum se známou hustotou pravděpodobnosti $p(v_k)$.

Cílem je určení podmíněné hustoty pravděpodobnosti $p(x_k \mid z^l)$.

# Výpočet aposteriorní filtrační hustoty

Předpokládejme, že bílé šumy $\{v_k\}, \{w_k\}$ jsou nezávislé a rovněž jsou nezávislé na $x_0$. Pak aposteriorní hustota pravděpodobnosti nebo alternativně filtrační hustota pravděpodobnosti může být určena rekurzivně

$$p(x_k \mid z^k) = \frac{p(z_k \mid x_k).p(x_k \mid z^{k-1})}{p(z_k \mid z^{k-1})}$$

kde $p(z_k \mid z^{k-1}) = \int p(z_k \mid x_k)p(x_k \mid z^{k-1})dx_k$.

### Výpočet prediktivní hustoty

Vývoj stavu je popsán

$$
\begin{aligned}
p(x_k \mid z^{k-1}) &= \int p(x_k, x_{k-1} \mid z^{k-1}) dx_{k-1} \\
&= \int p(x_k \mid x_{k-1}).p(x_{k-1} \mid z^{k-1}) dx_{k-1}
\end{aligned}
$$

přičemž $p(x_k \mid x_{k-1})$ je určeno $p(w_{k-1})$.

filtrační hustota

$$p(x_k \mid z^k) = \frac{p(x_k \mid z^{k-1}).p(z_k \mid x_k)}{p(z_k \mid z^{k-1})}$$

prediktivní hustota

$$p(x_k \mid z^{k-1}) = \int_{-\infty}^{\infty} p(x_{k-1} \mid z^{k-1}) p(x_k \mid x_{k-1}) dx_{k-1}$$

kde

$$p(z_k \mid z^{k-1}) = \int_{-\infty}^{\infty} p(x_k \mid z^{k-1}) p(z_k \mid x_k) dx_k$$

Uvažujme lineární gaussovský stochastický systém

$$x_{k+1} = F_k x_k + w_k \qquad k = 0, 1, 2, \ldots$$

$$z_k = H_k x_k + v_k \qquad k = 0, 1, 2, \ldots$$

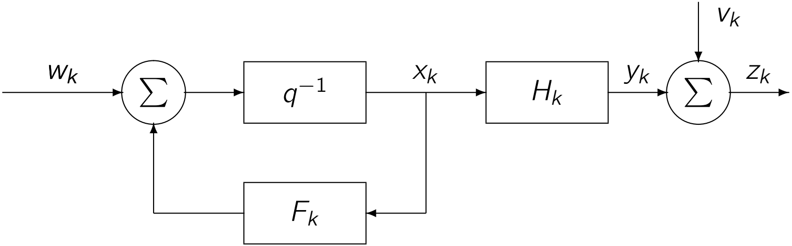
1. $\{v_k\}$ a $\{w_k\}$ jsou bílé šumy
2. $\{v_k\}$ a $\{w_k\}$ jsou gaussovské s nulovou střední hodnotou
3. $\{v_k\}$ a $\{w_k\}$ jsou nezávislé

$$E[v_k {v_l}^T] = R_k \delta_{k,l}$$

$$E[w_k {w_l}^T] = Q_k \delta_{k,l}$$

$$E[x_0] = \bar{x}_0 = \hat{x}_0' \qquad E[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T] = \Sigma_0 = P_0'.$$

# Lineární filtrace
## Lineární gaussovský systém




Signálový model lineárního systému.

- Pro aplikaci bayesovského přístupu potřebujeme rozhodnout, zda začít počítat filtrační nebo prediktivní hustotu.
- Přirozené je začít filtrací pro krok $k = 0$.
- Je nutné znát
  - $p(\mathbf{x}_k|\mathbf{z}^{k-1})$ - prediktivní hustota
  - $p(\mathbf{z}_k|\mathbf{x}_k)$ - hustota měření
  - $p(\mathbf{z}_k|\mathbf{z}^{k-1})$ - normalizační konstanta.

Prediktivní hustota

$$p(x_0 \mid z^{-1}) = N(x_0 : \hat{x}'_0, P'_0)$$

kde $E(x_0 \mid z^{-1}) = \hat{x}'_0$
$cov(x_0 \mid z^{-1}) = P'_0$.

$N(x_0 : \hat{x}'_0, P'_0)$ označuje, že náhodná veličina $x_0$ má normální rozložení se střední hodnotou $\hat{x}'_0$ a kovariancí $P'_0$.

Hustotu měření $p(z_k \mid x_k)$ snadno získáme

$$p(z_0 \mid x_0) = N(z_0 : H_0 x_0, R_0).$$

Normalizační konstantu

$$p(z_0 \mid z^{-1}) = N(z_0 : H_0 \hat{x}_0', H_0 P_0' H_0^T + R_0).$$

určíme takto

$$\begin{aligned} E[z_0 \mid z^{-1}] = E[H_0 x_0 + v_0 \mid z^{-1}] &= H_0 \hat{x}_0' \\ E[(z_0 - \hat{z}_0)(z_0 - \hat{z}_0)^T \mid z^{-1}] &= E[(H_0 x_0 + v_0 - H_0 \hat{x}_0') \\ &\quad (H_0 x_0 + v_0 - H_0 \hat{x}_0')^T \mid z^{-1}] \\ &= H_0 P_0' H_0^T + R_0. \end{aligned}$$

Nyní můžeme dosadit do Bayesova vztahu pro filtrační hustotu

$$
\begin{aligned}
p(x_0 \mid z^0) &= \frac{p(x_0 \mid z^{-1})p(z_0 \mid x_0)}{p(z_0 \mid z^{-1})} = \frac{N(z_0 : H_0 x_0, R_0)N(x_0 : \hat{x}'_0, P'_0)}{N(z_0 : H_0\hat{x}'_0, H_0 P'_0 H_0^T + R_0)} \\
&= \frac{1}{(2\pi)^{nz/2}(detR_0)^{1/2}} e^{-\frac{1}{2}[z_0 - H_0 x_0]^T R_0^{-1}[z_0 - H_0 x_0]} \\
&\times \frac{1}{(2\pi)^{nx/2}(detP'_0)^{1/2}} e^{-\frac{1}{2}[x_0 - \hat{x}'_0]^T P_0'^{-1}[x_0 - \hat{x}'_0]} \\
&\times \frac{(2\pi)^{nz/2}[det(H_0 P'_0 H_0^T + R_0)]^{1/2}}{e^{-\frac{1}{2}[z_0 - H\hat{x}'_0]^T [H_0 P'_0 H_0^T + R_0]^{-1}[z_0 - H\hat{x}'_0]}}.
\end{aligned}
$$

Po relativně zdlouhavých úpravách lze dokázat, že

$$p(x_0 \mid z_0) = N(x_0 : \hat{x}_0, P_0).$$

$$\hat{x}_0 = \hat{x}_0' + P_0' H_0^T [R_0 + H_0 P_0' H_0^T]^{-1} [z_0 - H_0 \hat{x}_0']$$
$$P_0 = P_0' - P_0' H_0^T [R_0 + H_0 P_0 H_0^T]^{-1} H_0 P_0'$$

kde

$$\hat{x}_0' = \mathrm{E}\left[x_0 | z^{-1}\right] = \bar{x}_0$$
$$P_0' = cov[x_0 | z^{-1}] = \Sigma_0.$$

Odvození prediktivní hustoty pravděpodobnosti $p(x_1 \mid z^0)$

přechodovou hustotu lze spočítat z modelu systému

$$p(x_1 \mid x_0) = N(x_1 : F_0 x_0, Q_0)$$

kde

$$\begin{aligned} E[x_1 \mid x_0] &= E[F_0 x_0 + w_0 \mid x_0] = F_0 x_0 \\ E[(x_1 - \hat{x}_1')((x_1 - \hat{x}_1')^T \mid x_0] &= E[(F_0 x_0 - F_0 x_0 + w_0) \\ &\quad (F_0 x_0 - F_0 x_0 + w_0)^T \mid x_0] \\ &= E[w_0 w_0^T \mid x_0] = Q_0. \end{aligned}$$

Dosazením do Bayesova vztahu

$$
\begin{aligned}
p(x_1 \mid z^0) &= \int N(x_0 : \hat{x}_0, P_0) N(x_1 : F_0 x_0, Q_0) dx_0 \\
&= \int \frac{1}{(2\pi)^{nx/2}.(detP_0)^{1/2}} e^{-\frac{1}{2}[x_0 - \hat{x}_0]^T P_0^{-1} [x_0 - \hat{x}_0]} \\
&\times \frac{1}{(2\pi)^{nx/2}(detQ_0)^{1/2}} e^{-\frac{1}{2}[x_1 - F_0 x_0]^T Q_0^{-1} [x_1 - F_0 x_0]} dx_0 \\
&= \frac{1}{(2\pi)^n x (detP_0)^{1/2}(detQ_0)^{1/2}} \\
&\times \int e^{-\frac{1}{2}\{[x_0 - \hat{x}_0]^T P_0^{-1}[x_0 - \hat{x}_0] + [x_1 - F_0 x_0]^T Q_0^{-1}[x_1 - F_0 x_0]\}} dx_0
\end{aligned}
$$

Po zdlouhavých úpravách dostaneme

$$p(x_1 \mid z^0) = N(x_1 : \hat{x}_1', P_1')$$

kde

$$\begin{aligned} \hat{x}_1' &= F_0 \hat{x}_0 \\ P_1' &= F_0 P_0 F_0^T + Q_0 \end{aligned}$$

$$\begin{aligned} \hat{x}_1' &= \mathrm{E}\left[x_1 | z^0\right] \\ P_1' &= cov[x_1 | z^0]. \end{aligned}$$

Výpočet filtrační hustoty pravděpodobnosti:

$$p(x_k \mid z^k) = N(x_k : \hat{x}_k, P_k) \qquad k = 0, 1, 2, \ldots$$

$$\hat{x}_k = \hat{x}'_k + P'_k H_k^T [R_k + H_k P'_k H_k^T]^{-1} [z_k - H_k \hat{x}'_k]$$

$$P_k = P'_k - P'_k H_k^T [R_k + H_k P_k H_k^T]^{-1} H_k P'_k.$$

Výpočet prediktivní hustoty pravděpodobnosti:

$$p(x_{k+1} \mid z^k) = N(x_{k+1} : \hat{x}'_{k+1}, P'_{k+1}) \quad k = 0, 1, 2, ..$$

$$\hat{x}'_{k+1} = F_k \hat{x}_k$$

$$P'_{k+1} = F_k P_k F_k^T + Q_k$$

# Lineární filtrace

## Kalmanův filtr ve formě prediktor

- Prediktor

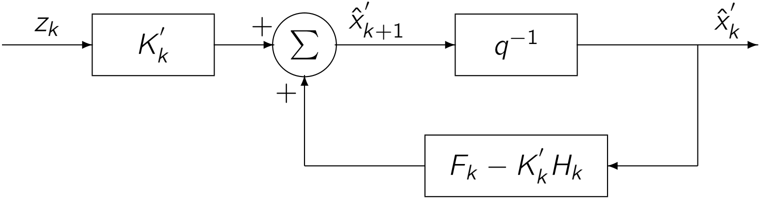
$$\hat{x}'_{k+1} = F_k[\hat{x}'_k + P'_k H_k^T (R_k + H_k P'_k H_k^T)^{-1}(z_k - H_k \hat{x}'_k)]$$
$$P'_{k+1} = F_k[P'_k - P'_k H_k^T (R_k + H_k P'_k H_k^T)^{-1} H_k P'_k] F_k^T + Q_k$$

-

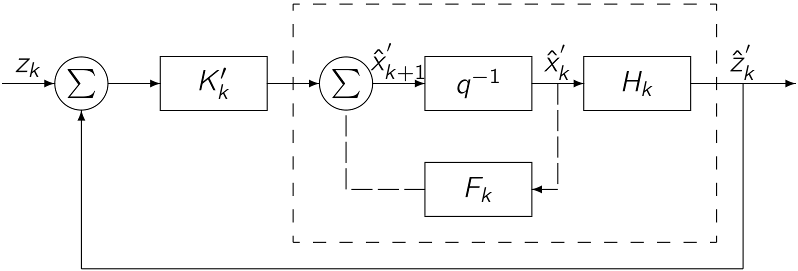
$$\hat{x}'_{k+1} = (F_k - K'_k H_k)\hat{x}'_k + K'_k z_k$$

- Kalmanovy zisky

$$K_k = P'_k H_k^T (R_k + H_k P'_k H_k^T)^{-1}$$
$$K'_k = F_k P'_k H_k^T (R_k + H_k P'_k H_k^T)^{-1}$$




Signálový model prediktoru




Signálový model prediktoru podle obsahující "kopii" struktury systému, jehož stav je odhadován

- Analytické řešení násobení normálních hustot

$$N(x_k : \hat{x}_k, P_k) = \frac{N(x_k : \hat{x}'_k, P'_k)N(z_k : H_k x_k, R_k)}{N(z_k : H_k \hat{x}'_k, H_k P'_k H_k^T + R_k)}$$

- Analytické řešení konvoluce

$$N(x_{k+1} : \hat{x}'_{k+1}, P'_{k+1}) = \int N(x_k : \hat{x}_k, P_k) N(x_{k+1} : F_k x_k, Q_k) dx_k$$

# Příklad 1:

Uvažujme skalární rovnici měření

$$z = x + v$$

$$\begin{aligned} p(x) &= \frac{1}{2} \qquad x \in (-1, 1) \\ &= 0 \qquad \text{jinak} \end{aligned}$$

$$\begin{aligned} p(v) &= \frac{1}{2} \qquad v \in (-1, 1) \\ &= 0 \qquad \text{jinak} \end{aligned}$$

Cílem je spočítat $p(x \mid z)$ a $E(x \mid z), cov(x \mid z)$.

Aproximační řešení náhradou $p(x)$ a $p(z)$ normálním rozložením

$$p_x(x) = \mathcal{N}\{x; 0, \tfrac{1}{3}\}$$

$$p_v(v) = \mathcal{N}\{v; 0, \tfrac{1}{3}\}$$

$$p(x|z) = \mathcal{N}\{x; \hat{x}, P\}$$

$$\hat{x} = \tfrac{1}{2}z$$

$$P = \tfrac{1}{6}$$

$$p(x \mid z) = \frac{p(z \mid x)p(x)}{p(z)} = \frac{p(z-x)p(x)}{\int p(z-x)p(x)dx}$$

kde

$$\begin{aligned} p(z-x) &= \frac{1}{2} \qquad z \in (-1+x, 1+x) \\ &= 0 \qquad \text{jinak} \end{aligned}$$

$$\begin{aligned} p(z) &= \frac{z+2}{4} \qquad z \in (-2, 0) \\ &= \frac{2-z}{4} \qquad z \in (0, 2) \\ &= 0 \qquad \text{jinak} \end{aligned}$$

$$p(x \mid z) = \frac{[sign(1+z-x) - sign(-1+z-x)][sign(x+1) - sign(x-1)]}{2[2 - zsign(z)][sign(z+2) - sign(z-2)]}$$

$$E(x \mid z) = z/2$$

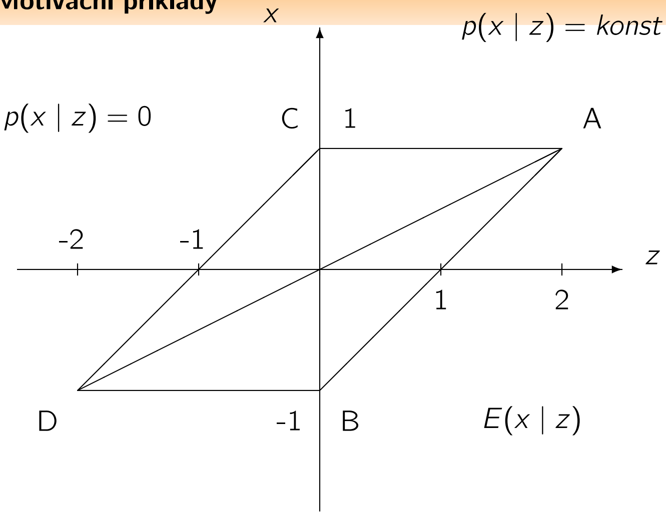
podmíněná kovariance pro $z \in (0, 2)$ je funkcí $z$

$$cov(x \mid z) = \frac{1}{3(2-z)}[1 - (z-1)^3] - \frac{z^2}{4}$$

pro $z \in (-2, 0)$

$$cov(x \mid z) = \frac{1}{3(z+2)}[1 + (z+1)^3] - \frac{z^2}{4}$$

opět funkcí $z$.




Podmíněná hustota pravděpodobnosti a střední hodnota

# Příklad 2:

Uvažujme skalární nelineární stavovou rovnici

$$x_{k+1} = f_k x_k + g_k x_k^2 + w_k$$

a dále předpokládejme

$$\begin{aligned} E(x_k \mid z^k) &= \hat{x}_k \\ E[(x_k - \hat{x}_k)^2 \mid z^k] &= P_k. \end{aligned}$$

Cílem je vypočítat

$$E(x_{k+1} \mid z^k) = \hat{x}'_{k+1} \quad a \quad cov(x_{k+1} \mid z^k) = P'_{k+1}.$$

Začneme střední hodnotou

$$
\begin{aligned}
\hat{x}'_{k+1} &= f_k E[x_k \mid z^k] + g_k E[x_k^2 \mid z^k] \\
&= f_k \hat{x}_k + g_k (P_k + \hat{x}_k^2)
\end{aligned}
$$

pro výpočet predikované střední hodnoty potřebujeme filtrační střední hodnotu a kovarianci

$$
\begin{aligned}
E(\tilde{x}_{k+1}^2 \mid z^k) &= (f_k + 2g_k\hat{x}_k)^2 E[\tilde{x}_k^2 \mid z^k) + g_k^2 E(\tilde{x}_k^4 \mid z^k) \\
&+ g_k^2 P_k^2 + E(w_k^2 \mid z^k) + 2g_k(f_k + 2g_k\hat{x}_k) E(\tilde{x}_k^3 \mid z^k) \\
&- 2g_k P_k (f_k + 2g_k\hat{x}_k) E(\tilde{x}_k \mid z^k) \\
&- 2g_k^2 P_k E(\tilde{x}_k^2 \mid z^k) = (f_k + 2g_k\hat{x}_k)^2 P_k + g_k^2 \gamma_k \\
&- g_k^2 P_k + Q_k + 2g_k (f_k + 2g_k\hat{x}_k)\delta_k
\end{aligned}
$$

kde $\gamma \triangleq E(\tilde{x}_k^4 \mid z^k)$ $\delta_k \triangleq E[\tilde{x}_k^3 \mid z^k]$.

- Lokální metody
  - aproximace nelineárních funkcí systému
  - výsledky mají lokální platnost
  - EKF, iterační filtry, filtr druhého řádu, bezderivační filtry (unscentovaná transformace, diference, ne derivace)
- Globální metody
  - pokrytí celého stavového prostoru
  - globální platnost

Uvažujme systém

$$x_{k+1} = f_k(x_k) + w_k$$

$$z_k = h_k(x_k) + v_k$$

$$p(x_0) = N(x_0 : \hat{x}'_0, P_0)$$

$$p(w_k) = N(w_k : 0, Q_k)$$

$$p(v_k) = N(v_k : 0, R_k)$$

Analytická řešitelnost filtračního kroku - diskuse

$$p(z_k \mid x_k) = N(z_k : h_k(x_k), R_k)$$

$$p(x_0 \mid z^0) = \frac{N(x_0 : \hat{x}_0', P_0')N(z_0 : h_0(x_0), R_0)}{\int N(x_0 : \hat{x}_0', P_0')N(z_0 : h_0(x_0), R_0)dx_0}$$

nejde analyticky řešit, protože $h(x_0)$ je nelineární funkce.

Aproximace nelineární funkce měření

$$h_0(x_0) \simeq h_0(\hat{x}_0') + H_0(\hat{x}_0')(x_0 - \hat{x}_0')$$

kde

$$H_0(\hat{x}_0') = \frac{\partial h_0(x_0)}{\partial x_0} \mid \hat{x}_0'.$$

Taylorův rozvoj $h_0(x_0)$ na okolí bodu $\hat{x}_0' = E(x_0 \mid z^{-1})$.

Bayesův vztah je analyticky řešitelný

$$p_A(x_0 \mid z^0) = \frac{N(x_0 : \hat{x}_0', P_0')N(z_0 : h_0(\hat{x}_0') + H_0(\hat{x}_0')(x_0 - \hat{x}_0'), R_0)}{\int N(x_0 : \hat{x}_0', P_0')N(z_0 : h_0(\hat{x}_0') + H_0(\hat{x}_0')(x_0 - \hat{x}_0'), R_0)dx_0}$$

protože

$$H_0(\hat{x}_0')x_0 + h_0(\hat{x}_0') - H_0(\hat{x}_0')\hat{x}_0'$$

je lineární funkce.

Aproximační filtrační hustota pravděpodobnosti

$$p_A(x_0 \mid z^0) = N(x_0 : \hat{x}_0, P_0)$$

$$\hat{x}_0 = \hat{x}'_0 + P'_0 H_0^T(\hat{x}'_0)[H_0(\hat{x}'_0)P'_0 H_0^T(\hat{x}'_0) + R_0]^{-1}[z_0 - h_0(\hat{x}'_0)]$$

$$P_0 = P'_0 - P'_0 H_0^T(\hat{x}'_0)[H_0(\hat{x}'_0)P'_0 H_0^T(\hat{x}'_0) + R_0]^{-1} H_0(\hat{x}'_0)P'_0$$

$H_0^T(\hat{x}'_0)$ je funkce měření.

Analytická řešitelnost prediktivního kroku - diskuse

$$p(x_1 \mid z^0) = \int p(x_0 \mid z^0) p(x_1 \mid x_0) dx_0$$

$$p(x_1 \mid x_0) = \mathcal{N}\{x_1 : f(x_0), Q_0\}$$

$$p_A(x_1 \mid z^0) = \int N(x_0 : \hat{x}_0, P_0) N(x_1 : f_0(x_0), Q_0) dx_0$$

nejde analyticky řešit, protože $f(x_0)$ je nelineární funkce.

Aproximace nelineární funkce dynamiky
Taylorův rozvoj funkce $f_0(x_0)$ na okolí $\hat{x}_0$

$$f_0(x_0) \simeq f_0(\hat{x}_0) + F_0(\hat{x}_0)(x_0 - \hat{x}_0)$$

$$F_0(\hat{x}_0) = \frac{\partial f(x_0)}{\partial x_0} \mid \hat{x}_0$$

Bayesův vztah analyticky řešitelný

$$p_A(x_1 \mid z^0) = \int N(x_0 : \hat{x}_0, P_0) N(x_1 : f_0(\hat{x}_0) + F_0(\hat{x}_0)(x_0 - \hat{x}_0), Q_0) dx_0$$

aproximační prediktivní hustota pravděpodobnosti

$$\begin{aligned} p_A(x_1 \mid z^0) &= N(x_1 : \hat{x}_1', P_1') \\ \hat{x}_1' &= f_0(\hat{x}_0) \\ P_1' &= F_0(\hat{x}_0') P_0 F_0^T(\hat{x}_0) + Q_0 \end{aligned}$$

$$F_0(\hat{x}_0)x_0 + f_0(\hat{x}_0) - F_0(\hat{x}_0)\hat{x}_0$$

je lineární funkce.

Rozšířený Kalmanův filtr (EKF)

Aproximační filtrační hustota pravděpodobnosti:

$$p_A(x_k \mid z^k) = N(x_k : \hat{x}_k, P_k)$$

$$\hat{x}_k = \hat{x}'_k + P'_k H_k^T(\hat{x}'_k)[H_k(\hat{x}'_k)P'_k H_k(\hat{x}'_k) + R_k]^{-1}[z_k - h_k(\hat{x}'_k)]$$

$$P_k = P'_k - P'_k H_k^T(\hat{x}'_k)[H_k(\hat{x}'_k)P'_k H_k(\hat{x}'_k) + R_k]^{-1} H_k(\hat{x}'_k)P'_k$$

Aproximační prediktivní hustota pravděpodobnosti:

$$\begin{aligned} p_A(x_{k+1} \mid z^k) &= N(x_{k+1} : \hat{x}'_{k+1}, P'_{k+1}) \\ \hat{x}'_{k+1} &= f_k(\hat{x}'_k) \\ P'_{k+1} &= F_k(\hat{x}_k)P_k F_k^T(\hat{x}_k) + Q_k \end{aligned}$$

Vlastnosti EKF

1. Kovarianční matice $P_k, P'_{k+1}$ není možné počítat off-line, protože jsou funkcemi měření
2. Filtrační a prediktivní hustoty jsou pouze aproximace skutečných hustot daných exaktním řešením problému na rozdíl od hustot generovaných Kalmanovým filtrem, které představují exaktní řešení problému.
3. Konvergence odhadu není zaručena
4. Oblíbený v praxi pro svoji jednoduchost

Aproximace Taylorovou řadou s ponecháním i členů druhého řádu

$$h_k(x_k) \simeq h_k(\hat{x}'_k) + H_k(\hat{x}'_k)(x_k - \hat{x}'_k) + \frac{1}{2}(x_k - \hat{x}'_k)^T \frac{\partial^2 h_k}{\partial x_k^2} \mid \hat{x}'_k (x_k - \hat{x}'_k)$$

kde

$$\frac{\partial^2 h_k}{\partial x_k^2} \mid \hat{x}'_k$$

je soubor $m$ matic $dim\ n \times n$ ($dim$ vektorové funkce $h_k(\cdot)$ je $m$ a $dim$ vektoru stavu $x_k$ je $n$).

$$h_k(\cdot) \triangleq [h_{1k}(\cdot), h_{2k}(\cdot), \cdots, h_{mk}(\cdot)]$$

$$M_{ik} \triangleq \quad \frac{\partial^2 h_{ik}}{\partial x_k^2} \mid \hat{x}'_k \quad = \quad \begin{bmatrix} \frac{\partial^2 h_{ik}}{\partial x_1 \partial x_1} & \frac{\partial^2 h_{ik}}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 h_{ik}}{\partial x_1 \partial x_n} \\ \vdots & & \vdots & \\ \frac{\partial^2 h_{ik}}{\partial x_n \partial x_1} & \ldots & ldots & \frac{\partial^2 h_{ik}}{\partial x_n \partial x_n} \end{bmatrix}$$

$$i = 1, 2, ..., m$$

$$\bar{h}_{ik} \triangleq (x_k - \hat{x}'_k)^T M_{ik} (x_k - \hat{x}'_k)$$

$$\bar{h}_k \triangleq \begin{bmatrix} \bar{h}_{1k} \\ \bar{h}_{2k} \\ \vdots \\ \bar{h}_{mk} \end{bmatrix}$$

$$h_k(x_k) \simeq h_k(\hat{x}'_k) + H_k(\hat{x}'_k)(x_k - \hat{x}'_k) + \bar{h}_k$$

Náhrada střední hodnotou - odstranění nelinearity

$$\bar{h}_{ik} = (x_k - \hat{x}'_k)^T M_{ik}(x_k - \hat{x}'_k) \simeq tr(P'_k M_{ik}) \stackrel{\triangle}{=} \bar{h}_{aik}$$

$$\bar{h}_{ak} = [\bar{h}_{a1k}, \bar{h}_{a2k}, \ldots, \bar{h}_{amk}]^T$$

$$h_k(x_k) \simeq h_k(\hat{x}'_k) + H_k(\hat{x}'_k)(x_k - \hat{x}'_k) + \bar{h}_{ak}$$

Filtrační hustota pravděpodobnosti

$$p_A(x_k \mid z^k) = N(x_k : \hat{x}'_k, P_k)$$

$$\hat{x}_k = \hat{x}'_k + P'_k H_k^T(\hat{x}'_k)[H_k(\hat{x}'_k)P'_k H_k^T(\hat{x}'_k) + R_k]^{-1}[z_k - h_k(\hat{x}'_k) - \bar{h}_{ak}]$$

$$P_k = P'_k - P'_k H_k^T(\hat{x}'_k)[H_k(\hat{x}'_k)P'_k H_k^T(\hat{x}'_k) + R_k]^{-1} H(\hat{x}'_k)P'_k$$

Rozvoj funkce $f_k(\cdot)$ pro predikci

$$f_k(x_k) \simeq f_k(\hat{x}_k) + F_k(\hat{x}_k)(x_k - \hat{x}_k) + \bar{f}_k.$$

kde

$$\bar{f}_k \triangleq \begin{bmatrix} \bar{f}_{1k} \\ \bar{f}_{2k} \\ \vdots \\ \bar{f}_{nk} \end{bmatrix}$$

# Základy nelineární filtrace

## Filtr druhého řádu

$$\bar{f}_{ik} \triangleq (x_k - \hat{x}_k)^T N_{ik} (x_k - \hat{x}_k)$$

$$N_{ik} \triangleq \frac{\partial^2 f_{ik}}{\partial x_k^2} \mid \hat{x}_k = \begin{bmatrix} \frac{\partial^2 f_{ik}}{\partial x_1 \partial x_1} & \frac{\partial^2 f_{ik}}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f_{ik}}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f_{ik}}{\partial x_2 \partial x_1} & \cdots & \cdots & \cdots \\ \vdots & & & \\ \frac{\partial^2 f_{ik}}{\partial x_n \partial x_1} & \cdots & \cdots & \frac{\partial^2 f_{ik}}{\partial x_n \partial x_n} \end{bmatrix}$$

$$i = 1, 2, \ldots, n$$

Odstraněním nelinearity ustředněním dostaneme

$$\bar{f}_{aik} = tr(P_k N_{ik})$$

$$\bar{f}_{ak} = [\bar{f}_{a1k}, \bar{f}_{a2k}, \ldots, \bar{f}_{ank}]^T$$

$$f_k(x_k) \simeq f_k(\hat{x}_k) + F_k(\hat{x}_k)(x_k - \hat{x}_k) + \bar{f}_{ak}$$

Prediktivní hustota pravděpodobnosti

$$p_a(x_{k+1} \mid z^k) = N(x_{k+1} : \hat{x}'_{k+1}, P'_{k+1})$$

$$\hat{x}'_{k+1} = F_k(\hat{x}_k) + \bar{f}_{ak}$$

$$P'_{k+1} = F_k(x_k) P_k F_k^T(x_k) + Q_k$$

Závěr: většinou kvalitnější odhad než EKF