

STATISTICKÉ ODHADY PARAMETRŮ

Jan Pech *

21. září 2001

1 Motivace

Obrazové snímače pracující ve vzdáleném infračerveném spektru jsou poměrně novou záležitostí. Ty nejkvalitnější snímače chlazené kapalným dusíkem jsou velmi drahé a jsou využívány převážně vojenským sektorem. Pro civilní využití se vyrábějí mnohem levnější snímače s menším rozlišením, schopné činnosti při pokojových teplotách.

Problémem těchto levných snímačů je však takzvaný *fixed pattern noise*, což je v podstatě rušivý obrazec sloučený se skutečným výstupním obrazem snímače. Tato prostorová nerovnoměrnost výstupního signálu je způsobena mikroskopickými rozdíly v rozměrech jednotlivých snímacích elementů a nepatrnými rozdíly teploty v ploše čipu snímače. Rušivý obrazec se navíc v čase pomalu mění, pravděpodobně díky teplotní závislosti. Rychlost a charakter těchto změn závisí na výrobní technologii obrazového snímače.

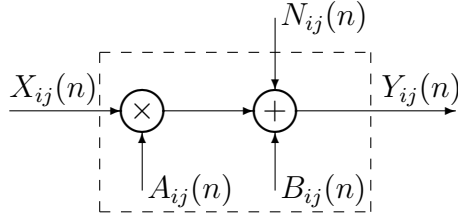
Výstupní signál jednoho snímacího elementu infračerveného obrazového snímače je možno pro nejjednodušší lineární model popsat vztahem:

$$Y_{ij}(n) = A_{ij}(n)X_{ij}(n) + B_{ij}(n) + N_{ij}(n), \quad (1)$$

kde $A_{ij}(n)$ je zisk (i, j) -tého detektoru v čase n , $B_{ij}(n)$ je jeho offset a $X_{ij}(n)$ je vstupní signál, v tomto případě intenzita záření. Člen $N_{ij}(n)$ reprezentuje aditivní šum o němž lze v prvním přiblížení předpokládat, že má normální rozdělení a je nekorelovaný se vstupním signálem. Graficky je tento model znázorněn blokovým schématem na obrázku 1.

V dalším textu budu předpokládat, že výstupní signál (i, j) -tého elementu $Y_{ij}(n)$ je v prostoru nekorelovaný s výstupními signály všech ostatních snímacích elementů $Y_{kl}(n); k \neq i, l \neq j$. Tak je možno pracovat se všemi výstupy jako s nezávislými jednorozměrnými signály. Měření potvrzují, že vazby mezi jednotlivými elementy jsou skutečně velice slabé, takže je možné alespoň v prvním přiblížení problém takto zjednodušit.

*ČVUT-FEL, K331, Technická 2, 166 27 Praha 6; pechj@fel.cvut.cz



Obrázek 1: Lineární model snímacího elementu

Zisk $A_{ij}(n)$ a offset $B_{ij}(n)$ (dále pouze $A(n)$ a $B(n)$) jsou ze statistického hlediska náhodné parametry. Pokud by byly známy jejich apriorní číselné charakteristiky nebo hustoty rozdělení, bylo by možné odhadovat jejich hodnoty. Při jiném úhlu pohledu můžeme označit vstupní signál $X_{ij}(n)$ (dále pouze $X(n)$) jako náhodný vektor. Při znalosti jeho číselných charakteristik nebo hustoty rozdělení je možné tento vektor odhadnout.

2 Obecné odhady

V dalším textu bude \mathbf{P} označovat vektor neznámých parametrů, které ovlivňují vektor měřených hodnot \mathbf{X} . Pokud je znám úplný statistický popis závislosti měřených hodnot na neznámých parametrech, je možné na základě naměřených hodnot \mathbf{x} vektoru \mathbf{X} určit odhad $\hat{\mathbf{P}}$ vektoru \mathbf{P} .

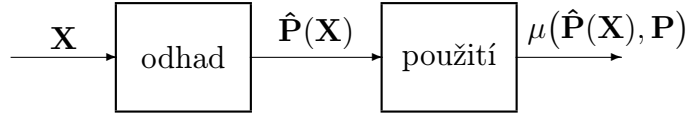
2.1 Odhad náhodných parametrů

Náhodné parametry mají charakter náhodné veličiny. To znamená, že známe jejich apriorní hustotu rozložení f_P , pokud nabývají spojitých hodnot nebo pravděpodobnostní funkci $Pr(\mathbf{P})$, pokud nabývají diskrétních hodnot. Další podmínkou je znalost úplného statistického popisu závislosti měřených hodnot na neznámých parametrech. Tento popis lze pro měřené hodnoty nabývající spojitých hodnot vyjádřit podmíněnou hustotou pravděpodobnosti $f_{X|P}$. Pro měřené hodnoty nabývající diskrétních hodnot lze tento popis vyjádřit podmíněnou pravděpodobností $Pr(\mathbf{X}|\mathbf{P})$.

Aby bylo možno stanovit kvalitu odhadu vektoru náhodných parametrů, je nutné stanovit jeho kritérium. Kritérium kvality odhadu však závisí na dalším využití daného odhadu, proto se pro posouzení kvality odhadu používají různá kritéria. Kvantitativní vyjádření (míra) následku chybného odhadu na kvalitu výsledného produktu se označuje $\mu(\hat{\mathbf{P}}, \mathbf{P})$ a je schematicky znázorněna na obrázku 2.

Nejllepší odhad je takový, který minimalizuje *riziko*, což je střední hodnota míry následku chybného odhadu:

$$E \left[\mu(\hat{\mathbf{P}}(\mathbf{X}), \mathbf{P}) \right] = \int_{\{\mathbf{X}\}} \int_{\{\mathbf{P}\}} \mu(\hat{\mathbf{P}}(\mathbf{x}), \mathbf{p}) f_{X|P}(\mathbf{x}, \mathbf{p}) d\mathbf{p} d\mathbf{x}. \quad (2)$$



Obrázek 2: Následek chyby odhadu

Úpravami uvedenými v [1] lze dostat obecný vzorec pro optimální odhad vektoru náhodných parametrů. Tento vzorec je však velice obecný a v praxi se využívají již odvozené vztahy pro některé běžné míry následku chybného odhadu. Dvě nejběžnější kritéria minimalizují středně kvadratickou hodnotu chyby odhadu nebo pravděpodobnost chyby pro diskrétní hodnoty.

2.1.1 MMSE

První kritérium minimalizuje středně kvadratickou hodnotu chyby. Zkratka MMSE znamená *minimum mean square error*. Odvození vzorce výpočtu MMSE odhadu vektoru parametrů je posáno v [1]. Výsledný vzorec pro měřené hodnoty nabývající spojitých hodnot bude:

$$\hat{\mathbf{P}}(\mathbf{X}) = \frac{\int_{\{\mathbf{P}\}} \mathbf{P} f_{X|P}(\mathbf{x}, \mathbf{P}) f_P(\mathbf{P}) d\mathbf{P}}{\int_{\{\mathbf{P}\}} f_{X|P}(\mathbf{x}, \mathbf{P}) f_P(\mathbf{P}) d\mathbf{P}} \quad (3)$$

Pro veličiny nabývající diskrétních hodnot je nutné do vzorce 3 dosadit vztah pro hustotu pravděpodobnosti této veličiny:

$$f_X(x) = \sum_{i=1}^n Pr(X = x_i) \delta(x - x_i) \quad (4)$$

MMSE odhad vektoru náhodných parametrů má následující vlastnosti:

- střední hodnota chyby odhadu je nulová (nestranný odhad)
- chyba je nekorelovaná s libovolnou funkcí měřených hodnot (ortogonalita)
- chyba odhadu je nekorelovaná s odhadem
- korelační matice chyby je rovna varianční matici chyby odhadu

2.1.2 MAP

Tento odhad minimalizuje pravděpodobnost chyby, je tedy nejpravděpodobnější. Zkratka MAP znamená *maximum a posteriori probability*. Míra následků chyby je pro toto kritérium dána:

$$\mu(\hat{P}, P) = \begin{cases} 0 & \text{pro } \hat{P} = P \\ 1 & \text{pro } \hat{P} \neq P \end{cases} \quad (5)$$

Toto kritérium je velmi výhodné například pro odhad dat, protože rizikem je pravděpodobnost záměny. Odhad minimalizující toto riziko poskytuje diskrétní hodnoty s minimální pravděpodobností chyby. Vzorec výpočtu odhadu vektoru náhodných parametrů podle kritéria MAP je opět odvozen v [1].

2.1.3 Odhad náhodného signálu

Pokud bude odhadovaný vektor náhodných parametrů závislý na čase, jedná se již o odhad náhodného signálu. Úkolem je odhadnout velikost náhodného signálu $\mathbf{p}(t)$ v čase t , přičemž máme k dispozici měřený signál $\mathbf{x}(\tau)$, $\tau \in (t_p, t_k)$. Podle vztahu mezi časem t a intervalem měření se tento odhad označuje jako:

- extrapolace: $(t < t_p) \cup (t > t_k)$
- interpolace: $t_p \leq t < t_k$
- filtrace: $t = t_k$

Při určování odhadu náhodného signálu se vychází z toho, že pro daný časový okamžik t je odhadovaný signál náhodným vektorem, takže je možné použít vztahy pro odhad náhodného vektoru parametrů. Do těchto vztahů je nutné dosadit:

$$\mathbf{P} = \mathbf{p}(t) \quad \Rightarrow \quad \hat{\mathbf{p}}(t) = \hat{\mathbf{P}}. \quad (6)$$

Pravděpodobnostní popis odhadovaného vektoru však závisí na zvoleném časovém okamžiku t , takže na něm bude záviset i algoritmus odhadu.

2.2 Odhad nenáhodných parametrů

Předchozí část vycházela z předpokladu znalosti apriorního rozložení neznámých parametrů. Tento předpoklad však v praxi často není splněn, takže uvedené vztahy není možné použít. Parametry u nichž neznáme apriorní rozložení nelze považovat za náhodné veličiny. K takovým parametrům je nutné přistupovat jako k nenáhodným veličinám o neznámé hodnotě. Skutečnost, že hustota pravděpodobnosti nebo pravděpodobnost závisí na vektoru neznámých parametrů \mathbf{P} se vyjádří zápisem $f_X(\mathbf{x}, \mathbf{P})$ nebo $Pr(\mathbf{X} = \mathbf{x}, \mathbf{P})$.

Odhady nenáhodných parametrů se v konkrétních případech řídí opět různými kritérii stejně jako tomu bylo u odhadů parametrů náhodných.

2.2.1 MVUB

Toto kritérium minimalizuje středně kvadratickou hodnotu chyby odhadu, která je rovna rozptylu odhadu. Zkratka MVUB znamená *minimum variance unbiased*. Odhad nenáhodného parametru podle tohoto kritéria je nestranný (nevychýlený), což

znamená že střední hodnota odhadu je rovna skutečné hodnotě parametru. Takovýto odhad nevykazuje systematickou chybu. Nevýhodou tohoto kritéria je, že nevychýlené odhady nemusí existovat. Zájemce o konkrétní vzorce odkazují přímo na skriptu [1], protože nalezení odhadu podle tohoto kritéria je velmi obtížné a pouhý výčet vzorců přesahuje rozsah tohoto dokumentu.

2.2.2 ML

Podle tohoto kritéria vybereme z možných hodnot nenáhodného parametru právě tu hodnotu, pro níž je nejvyšší hodnota podmíněné pravděpodobnosti $Pr(\mathbf{X} = \mathbf{x}, \hat{\mathbf{P}})$ pro diskrétní měřené hodnoty nebo podmíněné hustoty rozložení $f_X(\mathbf{x}, \hat{\mathbf{P}})$. Tento způsob odhadu se nazývá nejvěrohodnější. Zkratka ML znamená *maximum likelihood*.

Pokud bychom neznámé parametry považovali za náhodné s konstantní apriorní hustotou pravděpodobnosti, je ML odhad shodný s odhadem MAP náhodných vektorů.

2.2.3 LS

U tohoto kritéria uvažujeme, že vektor měřených hodnot \mathbf{X} nabývá hodnot v blízkosti nenáhodného vektoru $\mathbf{g}(\mathbf{P})$, který se nazývá model měřených hodnot. Odhad závisí pouze na volbě modelu \mathbf{g} , takže není nutné znát hustotu pravděpodobnosti měřených hodnot.

Mírou je součet kvadrátů odchylek jednotlivých hodnot:

$$\mu(\mathbf{X}, \mathbf{g}(\hat{\mathbf{P}})) = \sum_{i=1}^N (X_i - g_i(\hat{\mathbf{P}}))^2 \quad (7)$$

Odhad minimalizující tuto míru se nazývá odhad metodou nejmenších čtverců. Zkratka LS znamená *least squares*. Pro měření ve spojitém čase je model měřených hodnot funkce $h(t, \mathbf{P})$ a LS odhad je dán minimalizací výrazu:

$$\mu(\mathbf{X}, \mathbf{h}(\hat{\mathbf{P}})) = \int_L (x(t) - h(t, \hat{\mathbf{P}}))^2 dt \quad (8)$$

Tento odhad minimalizuje energii rozdílu měřeného signálu a jeho modelu. Odhad LS je shodný s odhadem ML pro aditivní šum s normálním rozložením o nulové střední hodnotě.

2.2.4 Momentová metoda

Protože odvození ML a LS odhadů je komplikované a vede obvykle k náročným algoritmům, používá se takzvaná momentová metoda. Tento postup sice nezaručuje optimalitu odhadu, ale lze jím odvodit odhady poměrně snadno.

Pro jednoduchost se pokusíme odhadnout jeden skalární parametr P z naměřených hodnot $x(n)$, $n = 0 \dots N-1$. Pokud mají všechny hodnoty $x(n)$ stejný obecný moment $\mu_i = E[x^i(n)]$, pak jej lze z naměřených hodnot odhadnout podle vztahu:

$$\hat{\mu}_i = \frac{1}{N} \sum_{n=0}^{N-1} x^i(n) \quad (9)$$

Protože hustota pravděpodobnosti naměřených hodnot závisí na odhadovaném parametru, závisí na něm i obecný moment $\mu_i = h_i(P)$. Odhad 9 je nestranný a rozptyl jeho chyby je nulový pro $N \rightarrow \infty$. Pro velký počet naměřených hodnot bude tedy platit $\hat{\mu}_i \rightarrow \mu_i = h_i(P)$. Parametr P lze tedy odhadnout podle vztahu:

$$\hat{P} = h_i^{-1} \left(\frac{1}{N} \sum_{n=0}^{N-1} x^i(n) \right), \quad (10)$$

kde h_i^{-1} je inverzní funkcí k funkci h_i . Pro velký počet měření se bude odhad blížit skutečné hodnotě parametru.

Momentovou metodou lze odvodit více odhadů v závislosti na tom, jaký řád i momentu zvolíme. Volba řádu je dána dvěma hledisky:

- řád by měl být co nejmenší, protože rozptyl odhadu s řádem roste.
- funkce h_i by měla být taková, aby inverzní funkce h_i^{-1} byla co nejjednodušší.

Pro odhad vektoru neznámých parametrů \mathbf{P} o k prvcích $P_1 \dots P_k$ pouze rozšíříme předchozí vztahy. Závislost momentů na tomto vektoru bude:

$$\begin{aligned} \mu_1 &= h_1(P_1 \dots P_k) \\ &\vdots \\ \mu_k &= h_k(P_1 \dots P_k) \end{aligned} \Leftrightarrow \boldsymbol{\mu} = \mathbf{h}(\mathbf{P}) \quad (11)$$

Odhad \mathbf{P} v závislosti na odhadech momentů bude:

$$\hat{\mathbf{P}} = \mathbf{h}^{-1}(\hat{\boldsymbol{\mu}}) = \mathbf{h}^{-1} \begin{bmatrix} \hat{\mu}_1 \\ \vdots \\ \hat{\mu}_k \end{bmatrix} = \mathbf{h}^{-1} \begin{bmatrix} \frac{1}{N} \sum_{n=0}^{N-1} x(n) \\ \vdots \\ \frac{1}{N} \sum_{n=0}^{N-1} x^k(n) \end{bmatrix} \quad (12)$$

2.2.5 Odhad nenáhodného signálu

Pokud odhadované neznámé parametry závisí na čase, jedná se již o odhad nenáhodného signálu $\mathbf{p}(t)$ v čase t . K dispozici máme pouze měřený signál $\mathbf{x}(\tau)$, $\tau \in (t_p, t_k)$. Stejně jako u odhadu náhodného signálu se tento odhad rozděluje na extrapolaci, interpolaci nebo filtraci podle vztahu mezi časem t a intervalem měření. Odhad pro konkrétní časový okamžik je možné provést podle některého z výše uvedených kritérií.

3 Lineární odhady

Lineární odhady jsou prakticky nejvýznamnějším případem suboptimálních odhadů náhodných parametrů. Tyto odhady lineárně závisí na naměřených hodnotách. Odhad se hledá ve tvaru:

$$\hat{\mathbf{P}}(\mathbf{X}) = \mathbf{a} + \mathbf{H}\mathbf{X}, \quad (13)$$

kde $\hat{\mathbf{P}}(\mathbf{X})$ je odhad neznámého vektoru a \mathbf{X} je vektor měřených hodnot. Vektor \mathbf{a} a matice \mathbf{H} jsou hledané koeficienty, nezávislé na odhadovaných parametrech ani na měřených hodnotách.

Narozdíl od obecných odhadů, kde se hledá optimální funkce, u lineárních odhadů se hledají pouze optimální hodnoty koeficientů. Díky tomu není nutné znát hustotu pravděpodobnosti naměřených hodnot, ale pouze jejich číselné charakteristiky.

3.1 Odhad náhodných parametrů

Pro lineární odhady vektorů náhodných parametrů \mathbf{P} je nutné znát jejich apriorní číselné charakteristiky podobně jako u obecných odhadů bylo nutné znát jejich apriorní hustotu rozložení.

3.1.1 Odhad náhodného vektoru

Jako kritérium kvality odhadu se zde uvažuje minimální středně kvadratická hodnota chyby odhadu. Odhad vektoru parametrů \mathbf{P} z vektoru měření \mathbf{X} je nutné určit tak, aby odhad vykazoval minimální korelační matici chyby odhadu. Postupem uvedeným v [1] dostaneme hodnoty koeficientů:

$$\mathbf{a} = \bar{\mathbf{P}} - \mathbf{H}\bar{\mathbf{X}} \quad (14)$$

$$\mathbf{H} = \text{cov}[\mathbf{P}, \mathbf{X}] \text{var}^{-1}[\mathbf{X}]. \quad (15)$$

Dosazením rovnice 14 do vztahu 13 určíme lineární odhad vektoru parametrů s nejmenší středně kvadratickou hodnotou chyby:

$$\hat{\mathbf{P}}(\mathbf{X}) = \bar{\mathbf{P}} + \mathbf{H}(\mathbf{X} - \bar{\mathbf{X}}), \quad (16)$$

kde matice \mathbf{H} je dána vztahem 15. K výpočtu odhadu z naměřených hodnot tedy stačí znalost následujících základních číselných charakteristik:

- střední hodnoty $\bar{\mathbf{P}}$ odhadovaného parametru,
- střední hodnoty $\bar{\mathbf{X}}$ měřených hodnot,
- varianční matice $\text{var}[\mathbf{X}]$ měřených hodnot,
- kovarianční matice $\text{cov}[\mathbf{P}, \mathbf{X}]$ odhadovaného parametru a měřených hodnot.

3.1.2 Odhad náhodného signálu

Pokud budou odhadované parametry záviset na čase, jedná se již o odhad velikosti náhodného signálu $\mathbf{p}(t)$ v čase t . K dispozici máme měřený signál $\mathbf{x}(\tau)$, $\tau \in (t_p, t_k)$. Stejně jako u obecných odhadů náhodného signálu se podle vztahu mezi časem t a intervalem měření rozlišuje extrapolace, interpolace a filtrace.

Při určování odhadu náhodného signálu se vychází z toho, že pro daný časový okamžik t je odhadovaný signál náhodným vektorem, takže je možné použít výše odvozené vztahy. Číselné charakteristiky odhadovaného vektoru však závisí na zvoleném časovém okamžiku t , takže na něm bude záviset i algoritmus odhadu.

3.2 Odhad nenáhodných parametrů

U nenáhodných parametrů neznáme jejich apriorní číselné charakteristiky, takže je musíme považovat pouze za nenáhodné parametry s neznámými hodnotami.

3.2.1 Odhad nenáhodného vektoru

Pokud bychom pro odhad nenáhodných parametrů použili vztah 16, dostaneme nepoužitelné řešení $\hat{\mathbf{P}} = \mathbf{P}$. Tento problém lze částečně řešit použitím nestranných odhadů.

Aby byl odhad popsán vztahem 13 nestranný, musí platit:

$$E[\hat{\mathbf{P}}] = \mathbf{a} + \mathbf{H}\bar{\mathbf{X}} \equiv \mathbf{P} \quad \forall \mathbf{P}, \quad (17)$$

což je zaručeno pokud je střední hodnota měřených veličin lineárně závislá na vektoru odhadovaných parametrů. Postupným odvozováním uvedeným v [1] se dá dostat až k určení koeficientů a z nich určit hodnotu odhadu nenáhodného vektoru.

3.2.2 Odhad nenáhodného signálu

Jak již bylo uvedeno výše, nenáhodný signál $\mathbf{p}(t)$ je známá funkce času $\mathbf{g}(\cdot)$, která je parametrizována neznámými koeficienty u_i , $i = 1 \dots N$:

$$\mathbf{p}(t) = \mathbf{g}(t, u_1 \dots u_N). \quad (18)$$

Teorii lineárního odhadu je možné aplikovat na náhodný signál ve speciálním tvaru:

$$\mathbf{p}(t) = \mathbf{A}(t)\mathbf{U} + \mathbf{b}(t), \quad (19)$$

kde $\mathbf{A}(t)$, $\mathbf{b}(t)$ jsou známé závislosti na čase a \mathbf{U} je vektor neznámých koeficientů. Pokud je totiž mezi dvěma vektory \mathbf{P} a \mathbf{U} lineární vztah, lze určit nejlepší nestranný lineární odhad vektoru \mathbf{P} z nejlepšího nestranného lineárního odhadu vektoru \mathbf{U} . Určíme tedy odhad vektoru koeficientů \mathbf{U} a z něho dostaneme odhad nenáhodného signálu:

$$\hat{\mathbf{p}}(t) = \mathbf{A}(t)\hat{\mathbf{U}} + \mathbf{b}(t). \quad (20)$$

4 Závěr

Účelem tohoto dokumentu není úplný výčet všech postupů odhadů náhodných i nenáhodných parametrů a odvození všech potřebných vztahů. Chtěl jsem pouze uvést krátký přehled možností statistického přístupu k odhadům náhodných signálů. Zájemce o tuto problematiku odkazuji na skripta [1], kde je tato tematika zpracována velmi podrobně, s odvozením všech vztahů a mnoha ilustračními příklady.

Literatura

- [1] Hrdina, Z.: *Statistická radiotechnika*. ČVUT, Praha 1996.
- [2] Kang, C.: *Computer Simulation of Compensation Method for Infrared Focal Plane Array*. Journal Beijing Institute of Technology, Vol. 9, No. 3, P. 330–334, Beijing 2000.
- [3] Harris, J., Chiang, Y.: *Nonuniformity Correction Using the Constant-Statistics Constraint: Analog and Digital Implementations*. SPIE Aerospace/Defense Sensing and Controls, Vol. 3061, 895-905, 1997.
- [4] Hayat, M., Torres, S.: *Statistical Algorithm for Nonuniformity Correction in Focal-Plane Arrays*. Applied Optics, Vol. 38, No. 8, 1999.