

APLIKACE INFORMAČNÍCH DIVERGENCÍ A PRAVDĚPODOBNOSTNÍCH METRIK V DEFEKTOSKOPII, ELASTICITĚ MATERIÁLŮ A ČÁSTICOVÉ FYZICE

ING. VÁCLAV KŮS, PH.D.

METRICKÉ VZDÁLENOSTI, DIVERGENCE, SKÓRY

Pro dvě distribuční funkce F, G na \mathbb{R} uvažujme např. následující typy vzdáleností a divergencí $D(F, G)$:

- Kolmogorovská vzdálenost: $K(F, G) = \sup_{x \in \mathbb{R}} |F(x) - G(x)|$
- Totální variace: $V(F, G) = \int |f - g| d\lambda$
- Cramér-von Mises vzd: $CM(F, G) = \int (F - G)^2 dF(x)$
- Lévy: $L(F, G) = \inf\{\varepsilon : G(x - \varepsilon) - \varepsilon \leq F(x) \leq G(x + \varepsilon) + \varepsilon, \forall x \in \mathbb{R}\}$
- Φ -divergence: $D_\phi(F, G) = \int g \phi(f/g) d\mu$, kde $\phi : (0, \infty) \rightarrow \mathbb{R}$, $\phi(1) = 0$, $\phi(t)$ konvexní na $(0, \infty)$ a striktně konvexní v $t = 1$, f a g jsou hustoty odpovídající F a G .
- Skórové M_ρ -funkce zobecňující ϕ -divergence

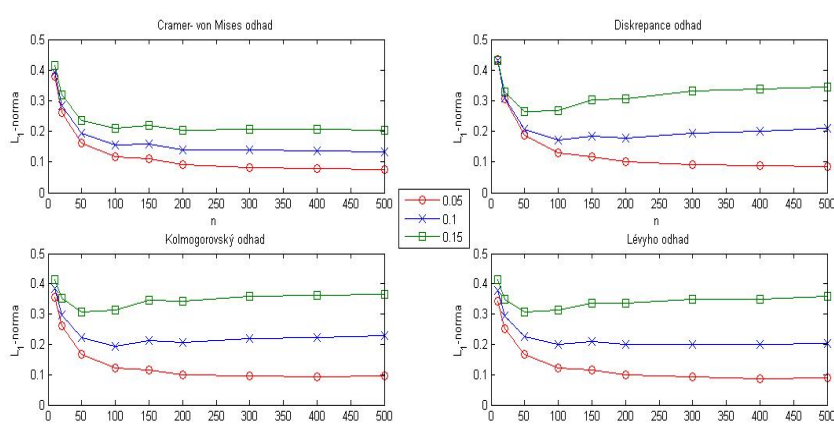
MDE ODHADY PRAVDĚPODOBNOSTNÍCH DISTRIBUCÍ

Pro soubor pozorování X_1, \dots, X_n definujeme empirickou distribuční funkci vztahem $F_n(x) = \frac{1}{n} \sum_{j=1}^n \mathbf{1}_{\{X_j \leq x\}}$. Pak za odhad skutečné hustoty pravděpodobnosti f_0 metodou minimální vzdálenosti (MDE) považujeme funkci \hat{f}_n , pro jejíž distribuční funkci \hat{F}_n platí: $D(\hat{F}_n, F_n) = \inf F D(F, F_n) + o(n^{-1/2})$, s.j., kde D je vybraná vzdálenost, divergence nebo skór. Říkáme, že \hat{f}_n je *konzistentní v ρ_D* vzdálenosti řádu $r_n \rightarrow 0$, pokud $\rho_D(\hat{f}_n, f_0) = O_p(r_n)$. Odhad považujeme za *robustní*, jestliže je založený na datovém vzorku ze znečištěné konvexní směsi $(1 - \varepsilon)f_0 + \varepsilon f_{kont}$ a přesto není výrazně ovlivněn kontaminující hustotou f_{kont} a silou znečištění ε .

Úkolem je vývoj a aplikace MDE odhadů:

- Konzistence, eficeince, kvalita pro neznečištěná data
- Monte-Carlo simulace, robustnost v závislosti na znečištění ε
- Aplikace na data z elastické defektoskopie materiálů a z fyziky elementárních částic (HEP)

GRAFY ROBUSTNOSTI VYBRANÝCH MDE ODHADŮ

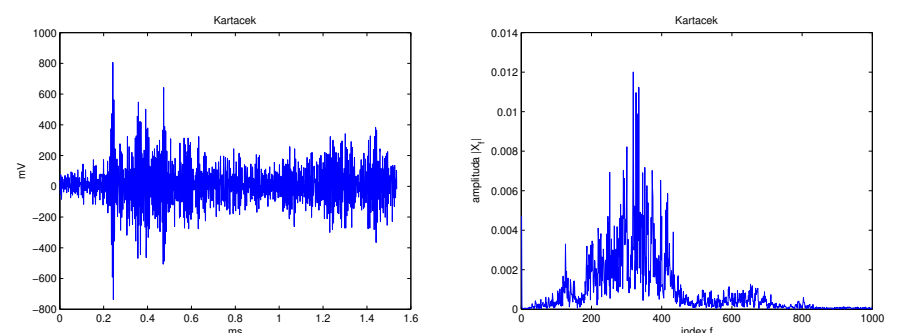


METODY KLASIFIKACE SIGNÁLŮ (STROJOVÉ UČENÍ)

Zdrojem klasifikovaných signálů jsou seismika, akustická emise defektů, koroze, valivých ložisek, tlakových nádob (jaderný reaktor), chemických zásobníků, přírodního potrubí, konstrukcí letadel, detektory elementárních částic (CERN, FERMILAB):

- Hierarchické a divizivní empirické metody založené na shlukování dat podle (ne)metrické míry podobnosti shluků
- Fuzzy klasifikátory založené na minimalizaci objektivní funkce s volitelným fuzzyfikačním faktorem
- Support Vector Machine (SVM) s lineárním, polynomiálním, exponenciálním jádrem
- EM odhady pravděpodobnostních hustot distribučních směrů (Model Based Clustering)
- Klasifikační regresní stromy/lesy s využitím metrik/divergencí
- Neuronové sítě, konvoluční a hluboké neuronové sítě
- Bayesovská klasifikace s použitím apriorní informace

SIGNÁL AKUSTICKÉ EMISE + SPEKTRÁLNÍ HUSTOTA



MBC KLASIFIKACE SIGNÁLŮ AKUSTICKÉ EMISE

Classification

