Úvod do neurovied

Informácie k zadaniu 1

Neurofyziologické dáta

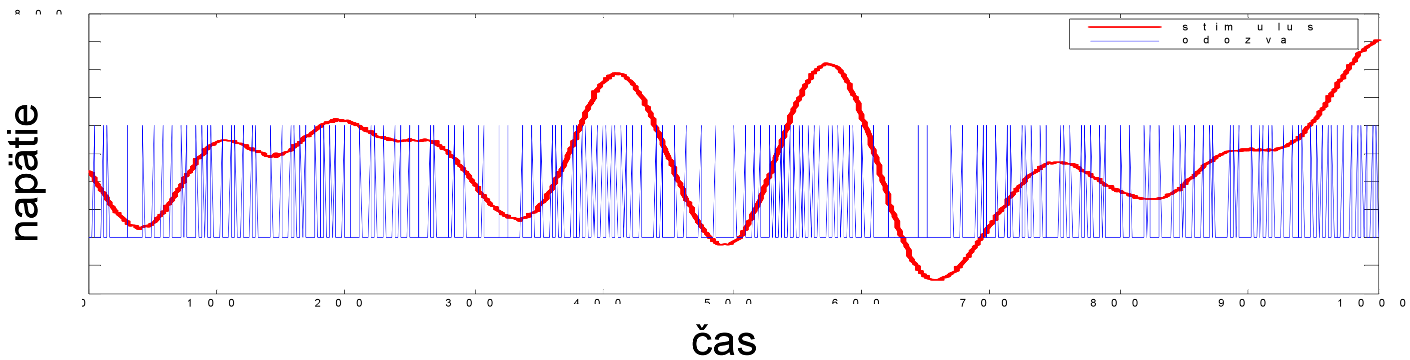
Prvé zadanie:

Ryba *Eigenmannia* (Wessel et al., 1996)

- má špeciálny orgán, ktorým generuje slabé oscilujúce elektrické pole s frekvenciou niekoľko sto Hz
- má elektrosenzorický orgán, citlivý na takéto elektrické pole
- používa ho pre elektrolokalizáciu a komunikáciu

Experiment:

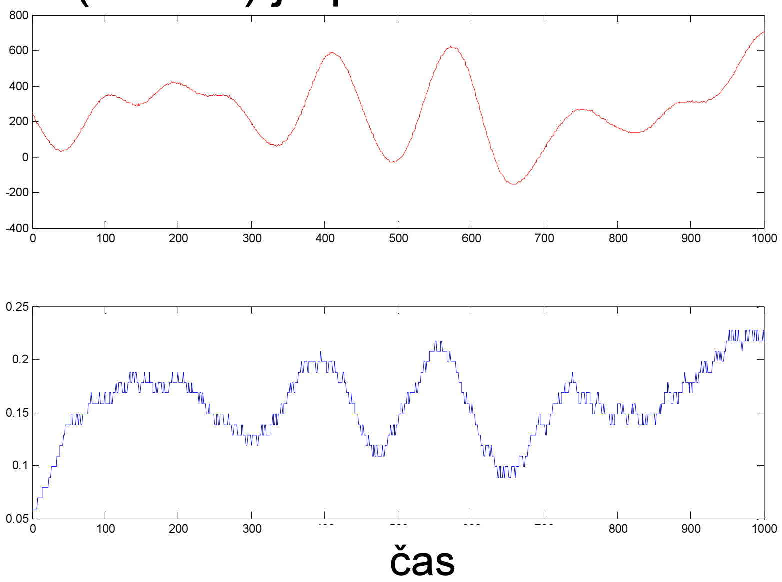
- umelo generované elektrické pole (amplitúdovo modulovaný sin signál) - červená
- zaznamenáva sa aktivita neurónu v elektrosenzor. orgáne – modrá
- Čo s tým???

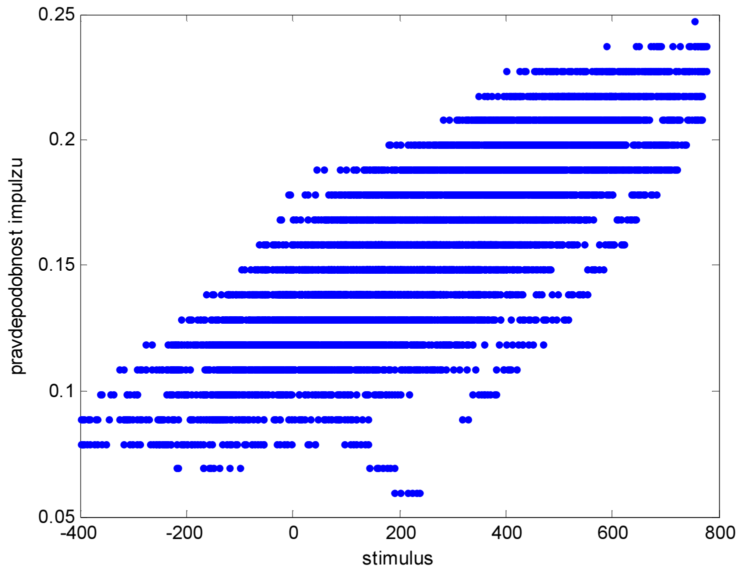


Neurofyziologické dáta – 1. zadanie

Tri vektory:

- čas (time, v milisek.), modulačný signál (stim), neurónové impulzy (rho)
- predošlý obrázok ukazuje signál a impulzy počas prvých 0.5 sek.
- pravdepodobnosť impulzu sa dá odvodiť konvolúciou s pravouhlým kernelom (viď neskôr)
- závislosť medzi signálom (červená) a pravdepodobnosťou impulzu (modrá) je približne lineárna





Neurofyziologické dáta – 1. zadanie

Takúto závislosť môžeme popísať lineárnym modelom:

$$y_i \approx a + bx_i$$

- kde *y* je pravdepodobnosť impulzu, *x* je stimul, *a* je priesečník priamky s osou Y, a *b* je smernica priamky
- *a, b* sú hľadané parametre modelu
- lineárny model je najpoužívanejším modelom vo všetkých disciplínach
- tento popis nebude bezchybný, dve hlavné príčiny: šum pri meraní a nelinearita dát
- štandardná metóda hľadania optimálneho *a* a *b* je metóda najmenších štvorcov, ktorá minimalizuje chybu:

$$E = \sum_{i=1}^{m} \frac{1}{2}(a + bx_i - y_i)^2$$

Neurofyziologické dáta – 1. zadanie

- Definovaním podmienky, že parciálne derivácie *E* podľa *a* a *b* musia byť rovné 0 vznikne systém dvoch lineárnych rovníc o dvoch neznámych

$$
\begin{aligned}
ma + b\sum_i x_i &= \sum_i y_i \\
a\sum_i x_i + b\sum_i x_i^2 &= \sum_i y_i x_i
\end{aligned}
$$

- substitúciou strednej hodnoty <x>, druhého momentu <x^2>, a korelácie <xy> sa rovnica zjednoduší na

$$\langle x\rangle = \frac{1}{m}\sum_{i=1}^{m} x_i \qquad \langle x^2\rangle = \frac{1}{m}\sum_{i=1}^{m} x_i^2$$

$$\langle y\rangle = \frac{1}{m}\sum_{i=1}^{m} y_i \qquad \langle xy\rangle = \frac{1}{m}\sum_{i=1}^{m} x_i y_i$$

$$
\begin{aligned}
a + b\langle x\rangle &= \langle y\rangle \\
a\langle x\rangle + b\langle x^2\rangle &= \langle xy\rangle
\end{aligned}
$$

- túto sústavu vyriešime vyjadrením *a* z prvej rovnice:

$$a = \langle y\rangle - b\langle x\rangle$$

Neurofyziologické dáta – 1. zadanie

- čím získame riešenie pre *b* v tvare:

$$b = \frac{\langle xy \rangle - \langle x \rangle \langle y \rangle}{\langle x^2 \rangle - \langle x \rangle^2}$$

- menovateľ tohto zlomku sa nazýva variancia (rozptyl) a udáva ako veľmi sa menia hodnoty *x*
- čitateľ sa nazýva kovariancia, a udáva vzájomnú závislosť medzi *x* a *y.*

- čiže na to, aby sme našli optimálny lineárny vzťah dvoch premenných, stačí nám vydeliť ich kovarianciu Cov(x,y) a varianciu Var(x).
- ďalší krok je už len dosadenie do rovnice pre výpočet *a*

- v MATLABe sa na určenie koeficientov optimálneho lineárneho modelu používa funkcia polyfit

Neurofyziologické dáta – 1. zadanie

- na vyhodnotenie lineárneho modelu sa používa korelačný koeficient *r*:

$$r = \frac{\langle xy \rangle - \langle x \rangle \langle y \rangle}{\sqrt{\langle x^2 \rangle - \langle x \rangle^2}\sqrt{\langle y^2 \rangle - \langle y \rangle^2}}$$

- *r* blízke ±1 hovorí, že lineárny model je dobrý, *r* blízke 0 hovorí, že nie
- *r* je vlastne kovariancia normalizovaná súčinom štandardných odchýliek
- často sa na hodnotenie lineárneho modelu používa r^2, ktoré hovorí, akú časť variancie v dátach je daný lineárny model schopný vysvetliť

Neurofyziologické dáta – vysvetlený zlomok variancie

Náš lineárny model je optimálny v zmysle najmenších štvorcov, t.j., chybu sme v ňom definovali ako varianciu odchýlky dát od priamky *ax*+*b*:

$$E = \sum_{i=1}^{m} \frac{1}{2}(a + bx_i - y_i)^2$$

Táto variancia (t.j., táto chyba) je nulová len keď je model dokonalý. Význam korelačného koeficientu sa dá chápať tak, že tento koeficient meria pokles vo variancii (tj, v chybe), ktorý získame použitím lineárneho modelu dát (model prvého rádu) oproti použitiu konštanty (čo by bol model nultého rádu). Aby sme našli najlepší model nultého rádu, stačí ak v hore uvedenej rovnici dáme b=0. E je v tomto prípade minimálne pre $a=\langle y \rangle$ (stredná hodnota y). Konkrétne bude hodnota E priamo úmerná variancii y.

Pri hľadaní najlepšieho modelu prvého rádu hľadáme hodnoty dvoch premenných a a b tak, aby sme minimalizovali E. Hodnota E sa v porovnaní s modelom nultého rádu zníži. A pomer tejto novej hodnoty E k tej starej je rovný $1-r^2$. To znamená, že r^2 vyjadruje zlomok variancie v y, ktorý bol vysvetlený pridaním lineárneho člena b do modelu.

Konvolúcia

- pre dva nekonečné časové rády, g_i a h_i, sa konvolúcia definuje ako

$$(g * h)_i = \sum_{j=-\infty}^{\infty} g_{i-j} h_j$$

- ak sú rády konečné, používa sa pri konvolúcii doplnenie nulami, alebo cyklické opakovanie rádov. Po doplnení nulami sa rovnica zmení na:

$$(g * h)_i = \sum_{j=0}^{n-1} g_{i-j} h_j$$

- aj keď rády *g* a *h* sú v rovnici rovnocenné, typicky sú ich funkcie odlišné: jednen z nich je zvyčajne dlhý, a reprezentuje signál, zatiaľčo druhý je krátky, koncentrovaný okolo 0, a nazýva sa filter alebo konvolučný kernel (jadro, maska)

Konvolúcia a pravdepodobnosť impulzov

- ak má g_i dĺžku *m* a h_i dĺžku *n*, potom výsledok konvolúcie má dĺžku *m+n-1*
- $(g*h)_0$ sa vypočíta ako

$$\begin{array}{ccccccccccc} \cdots & g_{m-1} & \cdots & g_1 & g_0 & 0 & 0 & 0 & 0 & \cdots \\ \cdots & 0 & \cdots & 0 & h_0 & h_1 & \cdots & h_{n-1} & 0 & \cdots \end{array}$$

- $(g*h)_1$ sa vypočíta ako

$$\begin{array}{ccccccccccc} \cdots & 0 & g_{m-1} & \cdots & g_1 & g_0 & 0 & 0 & 0 & \cdots \\ \cdots & 0 & \cdots & 0 & h_0 & h_1 & \cdots & h_{n-1} & 0 & \cdots \end{array}$$

- v MATLABe sa na výpočet konvolúcie používa funkcia `conv`
- na výpočet pravdepodobnosti impulzov zo zaznamenanej sekvencie impulzov sa používa konvolúcia v tvare:

$$p_i = \sum_j \rho_{i-j} w_j$$

- kde ρ je binárna sekvencia impulzov, *w* je jadro splňajúce podmienku, že suma(w)=1, a *p* je pravdepodobnosť impulzov. Najjednoduchšie jadro, ktoré spĺňa túto podmienku je pravoúhle jadro s dĺžkou *n* a hodnotou 1/*n*