Computational analysis of information-processing properties in an echo-state network.

Systémové rovnice

$$\tilde{x}(n) = tanh(W^{in}[u(n)] + Wx(n-1))$$
$$x(n) = (1-\alpha) * x(n-1) + \alpha * \tilde{x}(n)$$

kde

$$x(n) \in \mathbb{R}^{N_x}, \tilde{x}(n) \in \mathbb{R}^{N_x}$$
$$W^{in} \in \mathbb{R}^{N_x \times N_u}, W \in \mathbb{R}^{N_x \times N_x}$$
$$\alpha \in (0, 1]$$

Výstupná vrstva

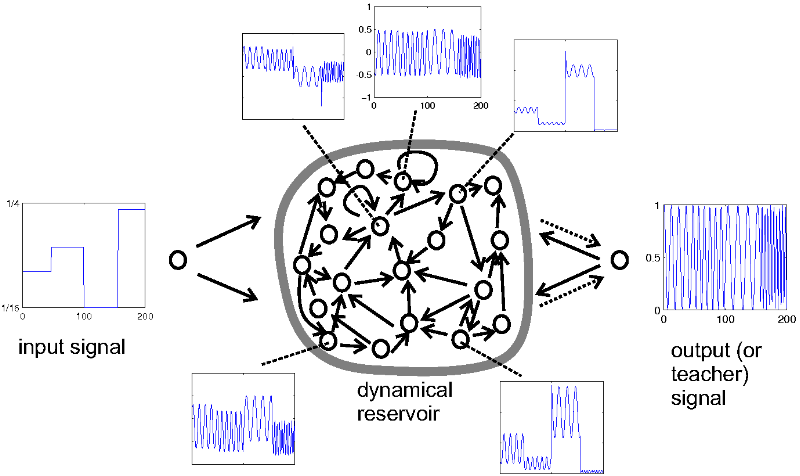
$$y(n) = W^{out}[x(n)]$$

kde

$$y(n) \in \mathbb{R}^{N_y}, W^{out} \in \mathbb{R}^{N_y \times (N_u + N_x)}$$

Sieť s echo stavmi

Model



Definícia: Predpokladajme, že vstupy pochádzajú z kompaktnej množiny U a že stavy siete ležia v kompaktnej množine A. Predpokladajme, že sieť nemá, žiadne spätnoväzobné spojenia. Potom má sieť vlastnosť echo stavu ak stav siete $x(n)$ je jednoznačne určený z ľava nekonečnou postupnosťou vstupov $\bar{u}^{-\infty}$. Presnejšie, pre každú postupnosť vstupov $\dots, u(n-1), u(n) \in U^{-\mathbb{N}}$, pre každú postupnosť stavov $\dots, x(n-1), x(n)$ a $\dots, x'(n-1), x(n) \in A^{-\mathbb{N}}$, kde $x(i) = T(x(i-1), u(i))$ a $x'(i) = T(x'(i-1), u(i))$ platí, že $x(n) = x'(n)$, kde T update operátor siete.

Sieť s echo stavmi
Krátkodobá pamäťová kapacita

Definícia: Nech $\nu(n) \in U$ (kde $-\infty < n < \infty$ a $U \subset \mathbb{R}$ je kompaktná) je jedno kanálový stacionárny vstupný signál. Predpokladajme, ďe máme rekureknťú neurónovú sieť špecifikovanú vnútornóu váhovou maticou $\mathbf{W}$, vstupným vektorom váh $\mathbf{w}^{in}$ a výstupnými funkciami neurónu f, f^{out}. Pre dané oneskorenie k a výstupný neurón y_k s výstupným vektorom spojení w_k^{out} uvažujeme koeficient determinácie

$$d[w_k^{out}](\nu(n-k), y_k(n)) = d(\nu(n-k), w_k^{out}\begin{pmatrix}\nu(n)\\ \mathbf{x}(n)\end{pmatrix}) =$$

$$= \frac{cov^2(\nu(n-k), y_k(n)))}{\sigma^2(\nu(n))\sigma^2(y_k(n))}$$

kde cov označuje kovarianciu a σ^2 varianciu.

1. *Krátkodobú pamäťovú kapacitu* siete k-teho oneskorenia definujeme

$$MC_k = \max_{w_k^{out}} d[w_k^{out}](\nu(n-k), y_k(n)).$$

2. *Krátkodobá pamäťová kapacita* siete je

$$MC = \sum_{k=1}^{\infty} MC_k$$

Základné pojmy informačnej teórie

Shannonova entropia

$$H(X) = -\sum_{x \in \Omega} p(x) \log p(x)$$

Podmienená entropia

$$H(X \mid Y) = -\sum_{x \in \Omega_x, y \in \Omega_y} p(x,y) \log p(x \mid y)$$

Vzájomná informácia

$$I(X : Y) = \sum_{x \in \Omega_x, y \in \Omega_y} p(x,y) \log \frac{p(x|y)}{p(x)p(y)} = H(X) - H(X \mid Y)$$

Podmienená vzájomná informácia

$$I(X : Y \mid Z) = \sum_{x \in \Omega_x, y \in \Omega_y, z \in \Omega_z} p(x,y,z) \log \frac{p(x \mid y,z)}{p(x \mid z)} =$$

$$= H(X \mid Z) - H(X \mid Y,Z)$$

Kozachenko - Leonenko estimator

$$H_{nn}(X) = \Psi(N) - \Psi(K) + \log c_d + \frac{d}{N}\sum_{i=1}^{N} \log \epsilon_i$$

kde $\Psi = \frac{d}{dx}\ln(\Gamma(x))$ (Digamma funkcia), $\frac{\epsilon_i}{2}$ je vzdialenosť K-teho najbližšieho susedav d-rozmernom priestore X, c_d je objem d-rozmernej jednotkovej gule.

Definicia

$$T^{(k,l)}_{Y \to X} = I(X_t : \mathbf{Y}^{(l)}_{t-u} \mid \mathbf{X}^{(k)}_{t-1})$$

kde

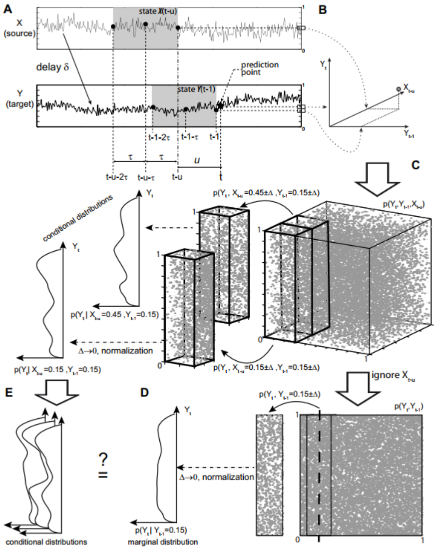
$$\mathbf{X}^{(k)}_{t-1} = (X_{t-1}, X_{t-1-\tau_k}, X_{t-1-2\tau_k}, \dots, X_{t-1-(k-1)\tau_k})$$
$$\mathbf{Y}^{(l)}_{t-u} = (Y_{t-u}, Y_{t-u-\tau_l}, Y_{t-u-2\tau_l}, \dots, Y_{t-u-(l-1)\tau_l})$$

Interpretácia

Transfer entropy je stupeň neurčitosti o súčasnom X vyriešený minulými X a Y, okrem stupňa neurčitosti o súčasnom X už vyriešenom minulosťou X samotného.

Transfer Entropy

Vizuálna sumarizácia



Kraskov-Grassberger-Stögbauer estimator

$$T^{(k,l)}_{Y\to X} = \Psi(K) + \langle \Psi(n_{\mathbf{x}^{(k)}_{t-1}}) - \Psi(n_{x_t, \mathbf{x}^{(k)}_{t-1}}) - \Psi(n_{\mathbf{x}^{(k)}_{t-1}, \mathbf{y}^{(l)}_{t-u}}) \rangle_{\mathrm{t}}$$

kde $\Psi = \frac{d}{dx} \ln(\Gamma(x))$ (Digamma funkcia), $\langle\cdot\rangle_t$ je priemer cez rôzne časové body, $n_{(\cdot)}$ je počet bodov vo vnútry hyperkociek okolo každého stavového vektora vo všetkých relevantných marginálnych priestoroch a polomer príslušných hyperkociek je daný K-tym najbližším susedom v najvyššorozmernom priestore generovanom $(x_t, \mathbf{x}^{(k)}_{t-1}, \mathbf{y}^{(l)}_{t-u})$.

Ragwitz-Kantz kritérium

Hľadáme

$$(k, \tau) = \underset{k \in \mathbb{Z}, \tau \in \mathbb{Z}}{\operatorname{argmin}} \|\hat{\mathbf{X}} - \mathbf{X}\|$$

kde

$$\mathbf{X} = (X_1, \ldots, X_n),\ \hat{\mathbf{X}} = (\hat{X}_1, \ldots, \hat{X}_n)$$

$$\hat{X}_{t+1} = \frac{1}{|\mathcal{U}_t|} \sum_{\vec{X}_l \in \mathcal{U}_t} X_{l+1}$$

$$\mathcal{U}_t = \{\vec{X}_l : \|\vec{X}_l - \vec{X}_t\| \leq \epsilon\},\ \vec{X}_t = (X_t, X_{t-\tau_k}, X_{t-2\tau_k}, \ldots, X_{t-(k-1)\tau}),$$

Permutačný test

H_0: medzi Y a X nie je žiaden informačný transfer.

P-hodnota:

$$P(T_{Y_{sur}\to X} \geq T_{Y\to X}) = \sum_{Y_{sur}:T_{Y_{sur}\to X}\geq T_{Y\to X}} P(Y_{sur})$$

kde Y_{sur} je permutácia $Y = (\mathbf{y}_{n-u}^{(l)}, \mathbf{y}_{n-1-u}^{(l)}, \ldots, \mathbf{y}_1^{(l)})$.

Active information storage

Definicia:

$$A_X^{(k)} = I(\mathbf{X}_{t-1}^{(k)} : X_t) = H(X_t) - H(X_t \mid \mathbf{X}_{t-1}^{(k)})$$

Interpretacia:

Koľko informácie je obsiahnutej v minulom stave $\mathbf{X}_{t-1}^{(k)}$ *premennej o jej ďalšom stave* X_t.

Kraskov-Grassberger-Stögbauer estimator:

$$A_X^{(k)} = \Psi(K) + \Psi(N) - \frac{1}{K} - \langle \Psi(n_{\mathbf{x}_{t-1}^{(k)}}) + \Psi(n_{x_t}) \rangle_t$$

Vzťah AIS a TE:

$$H(X_t) = A_X^{(k)}(t) + T_{Y \to X}^{(k,l)} + H(X_t \mid \mathbf{X}_{t-1}^{(k)}, \mathbf{Y}_{t-1}^{(l)})$$

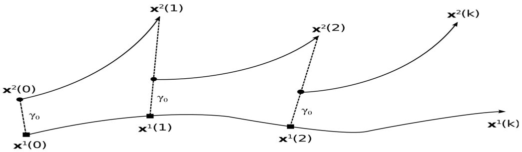
Lyapunov exponent

Definicia:

$$\lambda = \lim_{k\to\infty} \frac{1}{k} \ln(\frac{\gamma_k}{\gamma_0})$$

kde γ_0 je počiatočná vzdialenosť medzi perturbovanou a neperturbovanou trajektóriou dynamického systému, a γ_k je vzdialenosťv čase k. Pre sub-kritické systmémy $\lambda < 0$ a pre chaotické systémy $\lambda > 0$. Fázový prechod nastáva pri $\lambda = 0$ nazývaný tiež kritickým bodom, alebo hranicou chaosu.

Vizualizácia numerického odhadu Lyapunovho exponentu

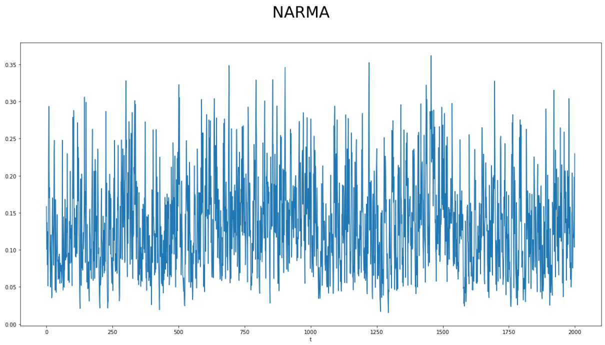


Benchmarkové časové rady

NARMA

$$y(t) = 0.2y(t) + 0.04y(t)\sum_{i=0}^{29} y(t-i) + 1.5x(t-29)x(t) + 0.001$$

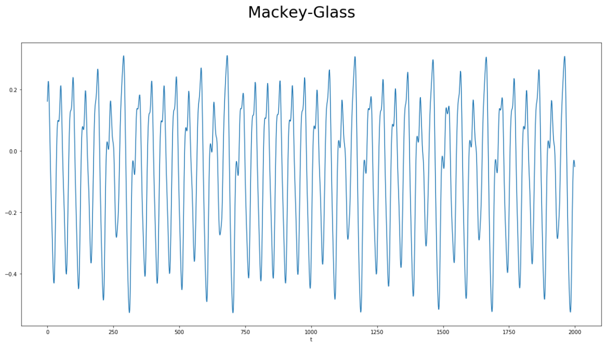
$$kde\ x(t) \sim unif(0, 0.5).$$



Benchmarkové časové rady
Mackey-Glass

$$\frac{dy}{dt} = 0.2\frac{y_{17}}{1 + y_{17}^{10}} - 0.1y$$

kde y_{17} je hodnota y v čase $t - 17$.



Benchmarkové časové rady
Lorenzov atraktor

$$\frac{dx}{dt} = 10(y - x)$$
$$\frac{dy}{dt} = x(28 - z) - y$$
$$\frac{dz}{dt} = xy - \frac{8}{3}z$$