Title: Empirical dynamic modeling

URL: https://en.wikipedia.org/wiki/Empirical_dynamic_modeling

PageID: 69664162

Categories: Category:Data modeling, Category:Machine learning, Category:Nonlinear systems, Category:Nonlinear time series analysis, Category:Predictive analytics

-----

Empirical dynamic modeling (EDM) is a framework for analysis and prediction of nonlinear dynamical systems . Applications include population dynamics , ecosystem service , medicine , neuroscience , dynamical systems , geophysics , and human-computer interaction . EDM was originally developed by Robert May and George Sugihara . It can be considered a methodology for data modeling , predictive analytics , dynamical system analysis, machine learning and time series analysis .

Description

Mathematical models have tremendous power to describe observations of real-world systems. They are routinely used to test hypothesis , explain mechanisms and predict future outcomes. However, real-world systems are often nonlinear and multidimensional, in some instances rendering explicit equation-based modeling problematic. Empirical models, which infer patterns and associations from the data instead of using hypothesized equations, represent a natural and flexible framework for modeling complex dynamics.

Donald DeAngelis and Simeon Yurek illustrated that canonical statistical models are ill-posed when applied to nonlinear dynamical systems. A hallmark of nonlinear dynamics is state-dependence: system states are related to previous states governing transition from one state to another. EDM operates in this space, the multidimensional state-space of system dynamics rather than on one-dimensional observational time series. EDM does not presume relationships among states, for example, a functional dependence, but projects future states from localised, neighboring states. EDM is thus a state-space , nearest-neighbors paradigm where system dynamics are inferred from states derived from observational time series. This provides a model-free representation of the system naturally encompassing nonlinear dynamics.

A cornerstone of EDM is recognition that time series observed from a dynamical system can be transformed into higher-dimensional state-spaces by time-delay embedding with Takens's theorem . The state-space models are evaluated based on in-sample fidelity to observations, conventionally with Pearson correlation between predictions and observations.

Methods

Primary EDM algorithms include Simplex projection, Sequential locally weighted global linear maps (S-Map) projection, Multivariate embedding in Simplex or S-Map, Convergent cross mapping (CCM), and Multiview Embeding, described below.

Nearest neighbors are found according to: $\text{NN}(y,X,k) = \|X_{N_i}^E - y\| \leq \|X_{N_j}^E - y\| \text{ if } 1 \leq i \leq j \leq k$

Simplex

Simplex projection is a nearest neighbor projection. It locates the k nearest neighbors to the location in the state-space from which a prediction is desired. To minimize the number of free parameters k is typically set to $E+1$ defining an $E+1$ dimensional simplex in the state-space. The prediction is computed as the average of the weighted phase-space simplex projected Tp points ahead. Each neighbor is weighted proportional to their distance to the projection origin vector in the state-space.

Find k nearest neighbor: $N_{k}\gets {\text{NN}}(y,X,k)$

Define the distance scale: $d\gets \|X_{N_{1}}^{E}-y\|$

Compute weights: For $i=1,\dots,k$: $w_{i}\gets \exp(-\|X_{N_{i}}^{E}-y\|/d)$

Average of state-space simplex: $\hat{y}\gets \sum_{i=1}^{k}\left(w_{i}X_{N_{i}+T_{p}}\right)/\sum_{i=1}^{k}w_{i}$

S-Map

S-Map extends the state-space prediction in Simplex from an average of the $E+1$ nearest neighbors to a linear regression fit to all neighbors, but localised with an exponential decay kernel. The exponential localisation function is $F(\theta)={\text{exp}}(-\theta d/D)$, where d is the neighbor distance and D the mean distance. In this way, depending on the value of θ, neighbors close to the prediction origin point have a higher weight than those further from it, such that a local linear approximation to the nonlinear system is reasonable. This localisation ability allows one to identify an optimal local scale, in-effect quantifying the degree of state dependence, and hence nonlinearity of the system.

Another feature of S-Map is that for a properly fit model, the regression coefficients between variables have been shown to approximate the gradient (directional derivative) of variables along the manifold. These Jacobians represent the time-varying interaction strengths between system variables.

Find k nearest neighbor: $N\gets {\text{NN}}(y,X,k)$

Sum of distances: $D\gets {\frac {1}{k}}\sum_{i=1}^{k}\|X_{N_{i}}^{E}-y\|$

Compute weights: For $i=1,\dots,k$: $w_{i}\gets \exp(-\theta \|X_{N_{i}}^{E}-y\|/D)$

Reweighting matrix: $W\gets {\text{diag}}(w_{i})$

Design matrix:
$$A\gets \begin{bmatrix}1&X_{N_{1}}&X_{N_{1}-1}&\dots &X_{N_{1}-E+1}\\1&X_{N_{2}}&X_{N_{2}-1}&\dots &X_{N_{2}-E+1}\\\vdots &\vdots &\vdots &\ddots &\vdots \\1&X_{N_{k}}&X_{N_{k}-1}&\dots &X_{N_{k}-E+1}\end{bmatrix}$$

Weighted design matrix: $A\gets WA$

Response vector at T_p :
$$b\gets \begin{bmatrix}X_{N_{1}+T_{p}}\\X_{N_{2}+T_{p}}\\\vdots \\X_{N_{k}+T_{p}}\end{bmatrix}$$

Weighted response vector: $b\gets Wb$

Least squares solution (SVD): $\hat{c}\gets {\text{argmin}}_{c}\|Ac-b\|_{2}^{2}$

Local linear model $\hat{c}$ is prediction: $\hat{y}\gets \hat{c}_{0}+\sum_{i=1}^{E}\hat{c}_{i}y_{i}$

Multivariate Embedding

Multivariate Embedding recognizes that time-delay embeddings are not the only valid state-space construction. In Simplex and S-Map one can generate a state-space from observational vectors, or time-delay embeddings of a single observational time series, or both.

Convergent Cross Mapping

Convergent cross mapping (CCM) leverages a corollary to the Generalized Takens Theorem that it should be possible to cross predict or cross map between variables observed from the same system. Suppose that in some dynamical system involving variables X {\displaystyle X} and Y {\displaystyle Y} , X {\displaystyle X} causes Y {\displaystyle Y} . Since X {\displaystyle X} and Y {\displaystyle Y} belong to the same dynamical system, their reconstructions (via embeddings) M_x {\displaystyle M_{x}} , and M_y {\displaystyle M_{y}} , also map to the same system.

The causal variable X {\displaystyle X} leaves a signature on the affected variable Y {\displaystyle Y} , and consequently, the reconstructed states based on Y {\displaystyle Y} can be used to cross predict values of X {\displaystyle X} . CCM leverages this property to infer causality by predicting X {\displaystyle X} using the M_y {\displaystyle M_{y}} library of points (or vice versa for the other direction of causality), while assessing improvements in cross map predictability as larger and larger random samplings of M_y {\displaystyle M_{y}} are used. If the prediction skill of X {\displaystyle X} increases and saturates as the entire M_y {\displaystyle M_{y}} is used, this provides evidence that X {\displaystyle X} is casually influencing Y {\displaystyle Y} .

Multiview Embedding

Multiview Embedding is a Dimensionality reduction technique where a large number of state-space time series vectors are combitorially assessed towards maximal model predictability.

Extensions

Extensions to EDM techniques include:

Generalized Theorems for Nonlinear State Space Reconstruction

Extended Convergent Cross Mapping

Dynamic stability

S-Map regularization

Visual analytics with EDM

Convergent Cross Sorting

Expert system with EDM hybrid

Sliding windows based on the extended convergent cross-mapping

Empirical Mode Modeling

Variable step sizes with bundle embedding

Multiview distance regularised S-map

See also

System dynamics

Complex dynamics

Nonlinear dimensionality reduction

References

Further reading

Chang, CW., Ushio, M. & Hsieh, Ch. (2017). "Empirical dynamic modeling for beginners" . Ecol Res . 32 (6): 785– 796. Bibcode : 2017EcoR...32..785C . doi : 10.1007/s11284-017-1469-9 . hdl : 2433/235326 . S2CID 4641225 . {{ cite journal }} : CS1 maint: multiple names: authors list (link)

Stephan B Munch, Antoine Brias, George Sugihara , Tanya L Rogers (2020). "Frequently asked questions about nonlinear dynamics and empirical dynamic modelling" . ICES Journal of Marine Science . 77 (4): 1463– 1479. doi : 10.1093/icesjms/fsz209 . {{ cite journal }} : CS1 maint: multiple names: authors list (link)

Donald L. DeAngelis, Simeon Yurek (2015). "Equation-free modeling unravels the behavior of complex ecological systems" . Proceedings of the National Academy of Sciences . 112 (13): 3856–3857. doi : 10.1073/pnas.1503154112 . PMC 4386356 . PMID 25829536 .

External links

State Space Reconstruction: Time Series and Dynamic Systems on YouTube

State Space Reconstruction: Takens' Theorem and Shadow Manifolds on YouTube

State Space Reconstruction: Convergent Cross Mapping on YouTube

EDM Introduction . Introduction with video, examples and references.

Geometrical theory of dynamical systems . Nils Berglund's lecture notes for a course at ETH at the advanced undergraduate level.

Arxiv preprint server has daily submissions of (non-refereed) manuscripts in dynamical systems.

Sugihara Lab , Scripps Institution of Oceanography, University of California San Diego.

Biological Nonlinear Dynamics Data Science Unit , Okinawa Institute of Science and Technology Graduate University, Okinawa Japan.