Time Series Analysis: Models and Methods (A Survey*)

*Some figures and sections are adopted from tutorial by E. Fox in ICML'15

The Data Science and Decisions Lab, UCLA

General Procedure for time series modelling: when, how and why?

Validation, statistical efficiency

Problem → **Model** → **Evaluation**

1- Forecasting
2- Hypothesis testing
3- Control
4- Clustering
5- Learning structure & dynamics

Pick a model based on the structure of the problem and the validity of assumptions

Medical Problems involving Time Series Analysis

1- Forecasting: -> Regression analysis
Predicting future values of physiological measurements, predicting length of a hospital’s waiting list, predicting population level spread of epidemics, survival analysis, etc.

2- Hypothesis testing: -> Early classification
Detecting disorders, confounding effects of treatments, testing patients’ latent classes, etc.

3- Control: -> Planning interventions over time

4- Clustering:-> Learning similarities across temporal data

5- Learning structure & dynamics:-> Learning causalities

Medical References involving Time Series Analysis

[1] Richman, Joshua S., and J. Randall Moorman. "Physiological time-series analysis using approximate entropy and sample entropy." *American Journal of Physiology-Heart and Circulatory Physiology* 278.6 (2000).

[2] Durbin, James, and Siem Jan Koopman. *Time series analysis by state space methods*. No. 38. Oxford University Press, 2012.

[3] Pincus, Steven M., and Ary L. Goldberger. "Physiological time-series analysis: what does regularity quantify?." *American Journal of Physiology-Heart and Circulatory Physiology* 266.4 (1994).

[4] Pope III, C. Arden, et al. "Respiratory health and PM10 pollution: a daily time series analysis." *American Review of Respiratory Disease* 144.3_pt_1 (1991).

[5] Wagner, Anita K., et al. "Segmented regression analysis of interrupted time series studies in medication use research." *Journal of clinical pharmacy and therapeutics* 27.4 (2002).

[6] Ivanov, Plamen Ch, et al. "Scaling behaviour of heartbeat intervals obtained by wavelet-based time-series analysis." *Nature* 383.6598 (1996): 323-327.

[7] van Walraven, Carl, Vivek Goel, and Ben Chan. "Effect of population-based interventions on laboratory utilization: a time-series analysis." *Jama* 280.23 (1998): 2028-2033.

[8] Friston, Karl J., Peter Jezzard, and Robert Turner. "Analysis of functional MRI time-series." *Human brain mapping* 1.2 (1994): 153-171.

[9] Kantelhardt, Jan W., et al. "Multifractal detrended fluctuation analysis of nonstationary time series." *Physica A: Statistical Mechanics and its Applications* 316.1 (2002): 87-114.

[10] Fritsche, Lutz, et al. "Recognition of critical situations from time series of laboratory results by case-based reasoning." *Journal of the American Medical Informatics Association* 9.5 (2002): 520-528.

[11] Perotte, Adler, et al. "Risk prediction for chronic kidney disease progression using heterogeneous electronic health record data and time series analysis."*Journal of the American Medical Informatics Association* 22.4 (2015): 872-880.

Medical References involving Time Series Analysis

[12] Dowding, Dawn W., Marianne Turley, and Terhilda Garrido. "The impact of an electronic health record on nurse sensitive patient outcomes: an interrupted time series analysis." *Journal of the American Medical Informatics Association* 19.4 (2012): 615-620.

[13] Cheng, Karen Elizabeth, et al. "Structural models used in real-time biosurveillance outbreak detection and outbreak curve isolation from noisy background morbidity levels." *Journal of the American Medical Informatics Association* 20.3 (2013): 435-440.

[14] Paxton, Chris, Alexandru Niculescu-Mizil, and Suchi Saria. "Developing predictive models using electronic medical records: challenges and pitfalls."*AMIA Annual Symposium proceedings/AMIA Symposium. AMIA Symposium*. Vol. 2013. American Medical Informatics Association, 2012.

[15] Dyagilev, Kirill, and Suchi Saria. "Learning (predictive) risk scores in the presence of censoring due to interventions." *Machine Learning* (2015): 1-26.

[16] Wang, Yuan, et al. "Mortality Prediction in ICUs Using A Novel Time-Slicing Cox Regression Method." *AMIA Annual Symposium Proceedings*. Vol. 2015. American Medical Informatics Association, 2015.

[17] Caballero, Karla, and Ram Akella. "Dynamic Estimation of the Probability of Patient Readmission to the ICU using Electronic Medical Records." *AMIA Annual Symposium Proceedings*. Vol. 2015. American Medical Informatics Association, 2015.

[18] Stanculescu, Ioan, Christopher KI Williams, and Yvonne Freer. "Autoregressive hidden Markov models for the early detection of neonatal sepsis." *Biomedical and Health Informatics, IEEE Journal of* 18.5 (2014): 1560-1570.

[19] Henry, Katharine E., et al. "A targeted real-time early warning score (TREWScore) for septic shock." *Science Translational Medicine* 7.299 (2015): 122-299.

[20] Ghassemi, Marzyeh, et al. "A Multivariate Timeseries Modeling Approach to Severity of Illness Assessment and Forecasting in ICU with Sparse, Heterogeneous Clinical Data." *AAAI*. 2015.

… and many more…

How to start?

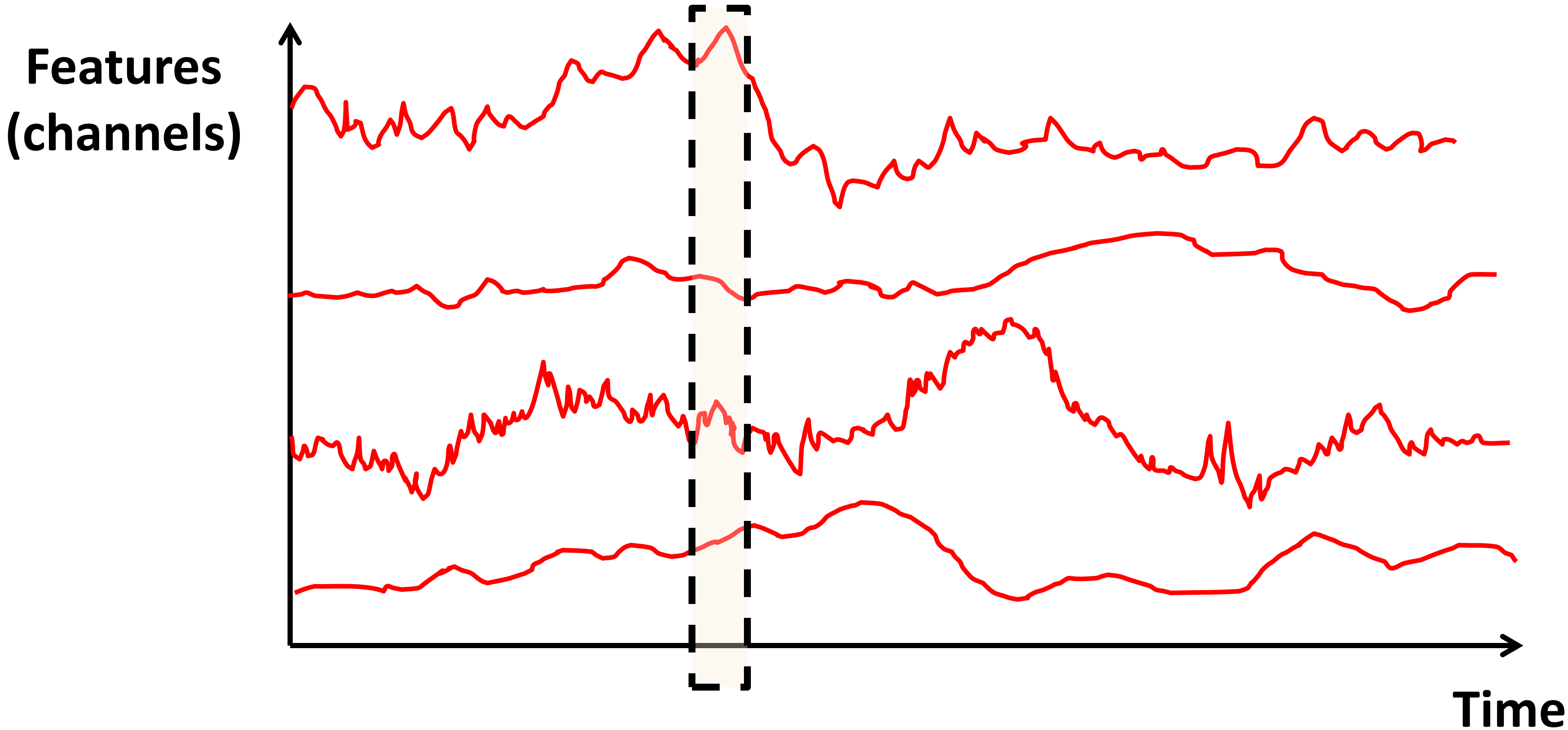
Intuition and understanding the idiosyncrasies of the problem!

- Is the series stationary? Disorders?
- Strong mixing?
- Multiple dependent series?
- Causalities between series?
- Exogenous interventions?

→ Pick a set of Models

⇅ Evaluate and eliminate

The two dimensions of multivariate time series



Every model captures:

Evolution: **impacts timeliness of decisions and accuracy of predictions**

Relational structure: **impacts cost, computational and statistical efficiency**

1) Autoregressive Moving Average Models (ARMA)

Assumptions

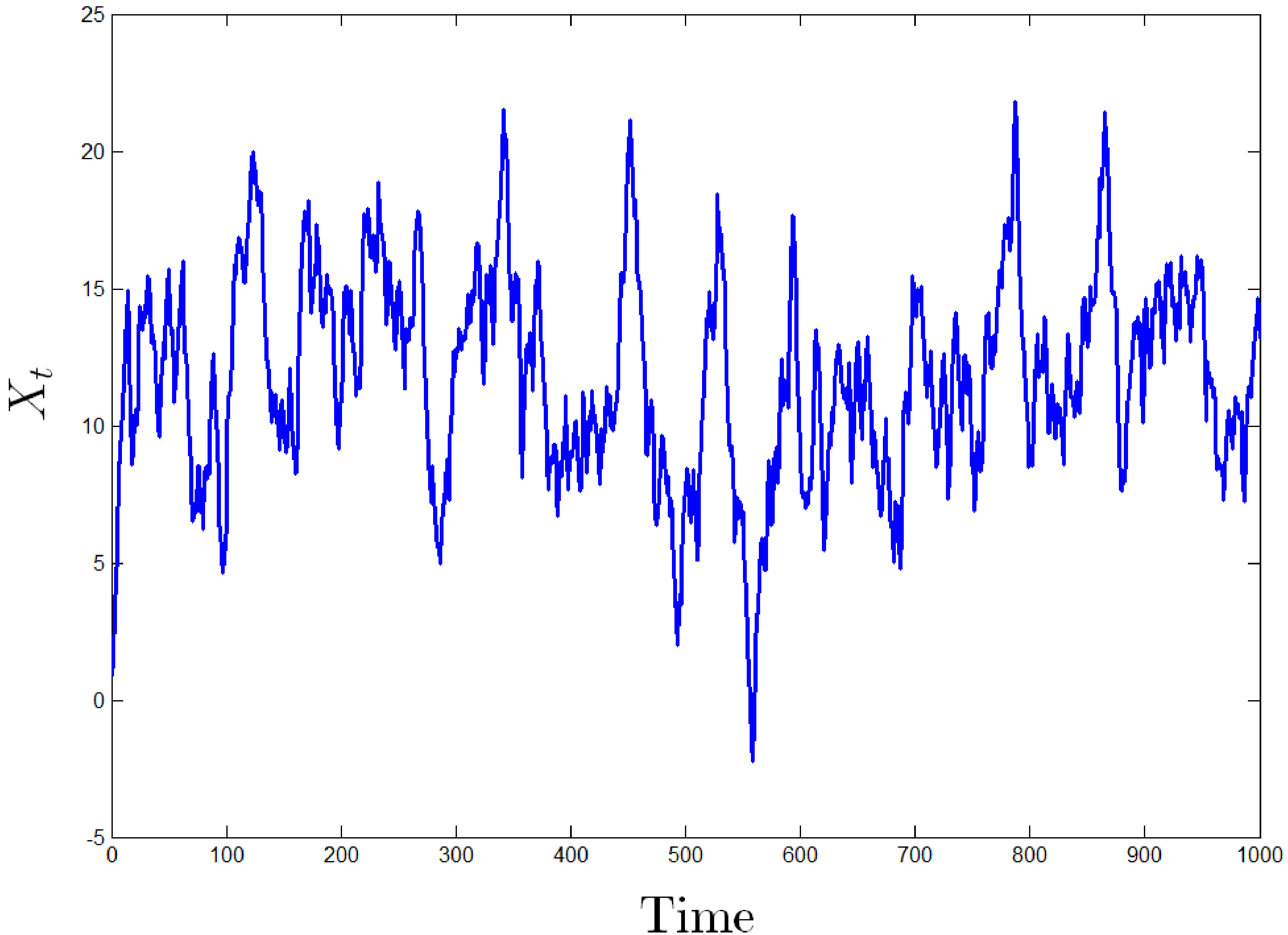
Underlying process being modeled is (weakly) stationary: the process exhibits no trends over time (i.e. constant mean and variance).

Construction: univariate ARMA(p,q) model

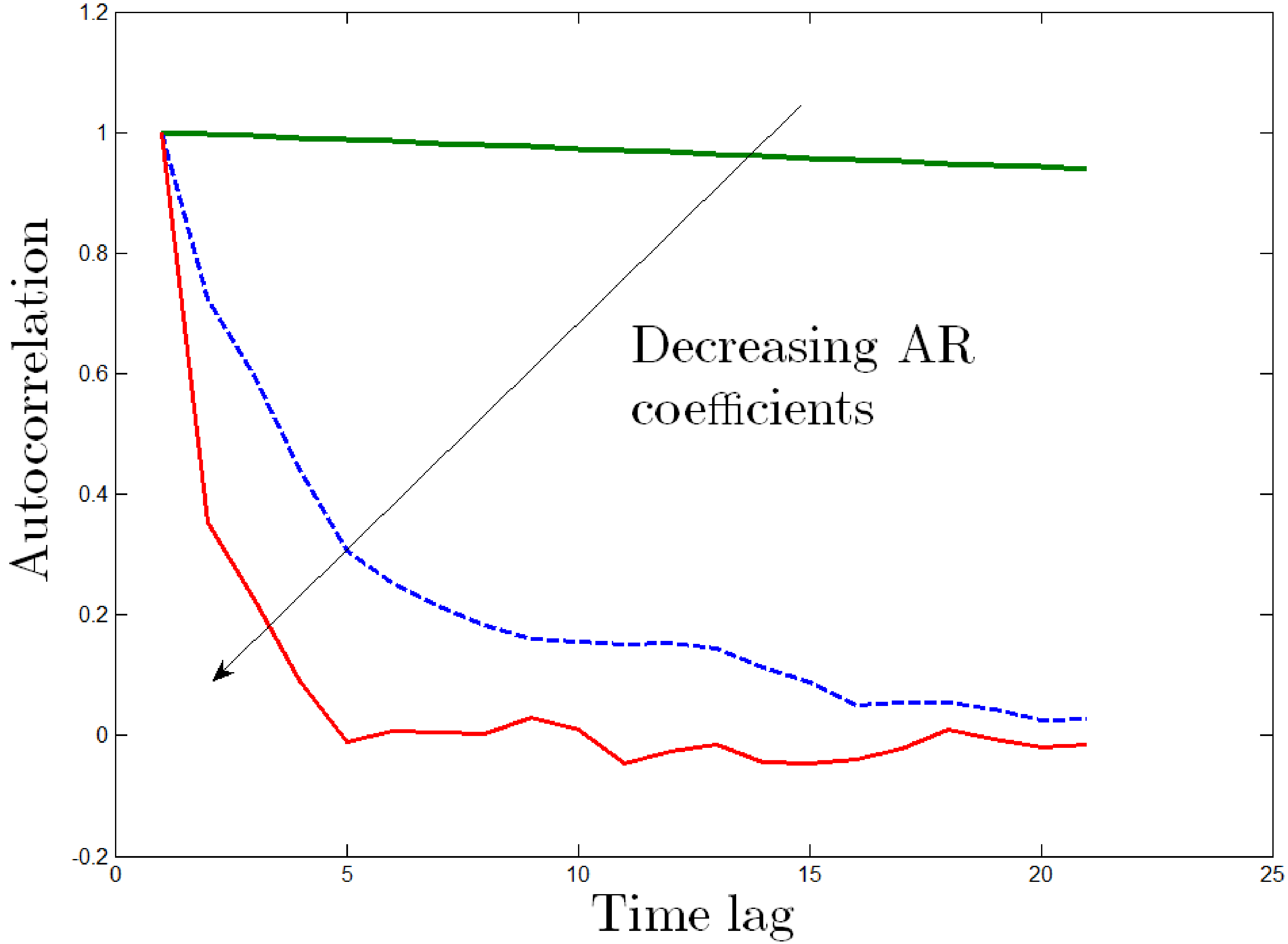
$$X_t = c + \epsilon_t + \sum_{i=1}^{p} \alpha_i \, X_{t-i} + \sum_{i=1}^{q} \beta_i \, \epsilon_{t-i}$$

- c: **Constant**
- ϵ_t: **Gaussian noise**
- α_i: **AR coefficients**
- β_i: **MA coefficients**

1) ARMA: stationarity of the generative model



1) ARMA model capturing evolution



1) Multivariate ARMA model

Evolution described by AR and MA coefficients, what about relational structure?

Construction: multivariate ARMA(p,q) model

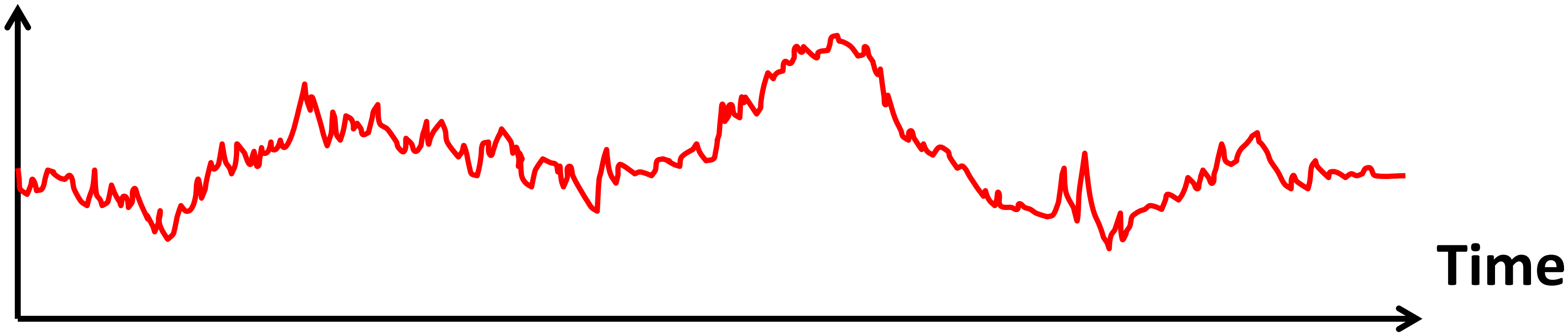
$$\mathbf{X}_t = c + \epsilon_t + \sum_{i=1}^{p} \mathbf{A}_i \, \mathbf{X}_{t-i} + \sum_{i=1}^{q} \mathbf{B}_i \, \epsilon_{t-i}$$

Companion matrix (AR matrix) and MA matrix capture the relational structure

Variant constructions can capture drifts, periodicity, etc. E.g. ARIMA and seasonal ARIMA.

1) Fitting the ARMA model

Learning the parameters of the ARMA(p,q) model (model fitting)



Model fitting = find AR, MA coefficients, and number of such coefficients!

$$X_t = c + \epsilon_t + \sum_{i=1}^{p} \alpha_i X_{t-i} + \sum_{i=1}^{q} \beta_i \epsilon_{t-i}$$

1) Fitting the ARMA model

Learning the parameters of the ARMA(p,q) model (model fitting)

1) **Estimate the number of coefficients p and q**

Use an information criterion to select a model (e.g. Akaike IC and Bayesian IC).

2) **Maximum Likelihood Estimation**

Estimating the covariance and mean parameters as a function of the AR and MA parameters.

1) Fitting the ARMA model

Learning the parameters of the ARMA(p,q) model (model fitting)

1) Estimate the number of coefficients p and q: use AIC

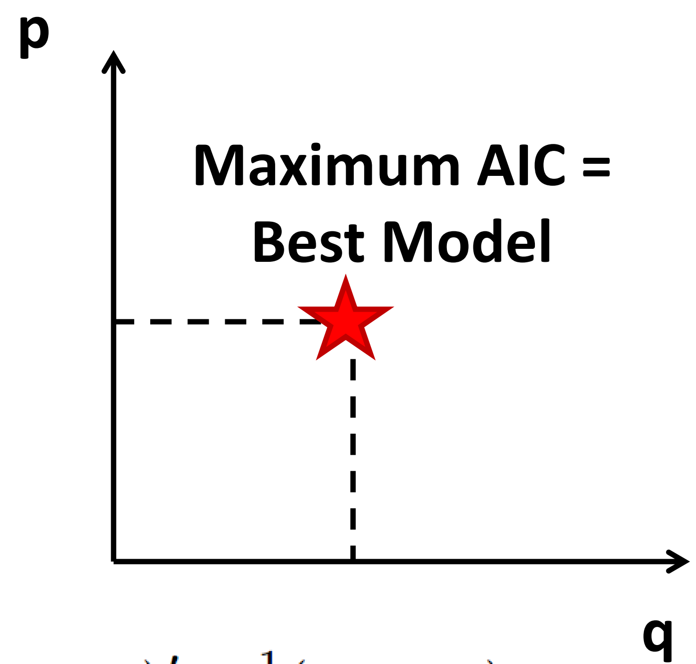
- Measures relative quality of statistical models for a given set of data.
- Relative estimate of the information lost when a given model is used to represent the process that generates the data.
- Trade-off between the goodness-of-fit of the model and the complexity of that model.

1) Fitting the ARMA model

Learning the parameters of the ARMA(p,q) model (model fitting)

Estimate the number of coefficients p and q: use AIC

$$\mathrm{AIC} = 2pq - 2\log\left(\ell(p,q)\right)$$



MLE for specific p and q

$$-2\ell(\mu,\phi,\theta,\sigma^2) = n\log 2\pi + \log|\Gamma_n| + (\boldsymbol{X}-\mu)'\Gamma_n^{-1}(\boldsymbol{X}-\mu)$$

1) ARMA model: pros and cons

Advantages

Stationarity assumption (coefficients are constant over time) leads to statistical efficiency and ease of construction

Limitations

Cannot capture complicated or non-stationary dynamics unless upgraded in a way that sacrifices statistical efficiency

ARMA is usually used in quantitative finance, econometrics, weather forecasting, etc, in order to predict future values of a series

1) ARMA model: key papers

[1] Harrison, L., William D. Penny, and Karl Friston. "Multivariate autoregressive modeling of fMRI time series." *NeuroImage* 19.4 (2003): 1477-1491.

[2] Weber, Edith JM, Peter Molenaar, and Maurits W. Molen. "A nonstationarity test for the spectral analysis of physiological time series with an application to respiratory sinus arrhythmia." *Psychophysiology* 29.1 (1992): 55-62.

[3] Kis, Maria. "Time Series Models on Analysing Mortality Rates and Acute Childhood Lymphoid." *Connecting Medical Informatics and Bio-informatics: Proceedings of MIE2005: the XIXth International Congress of the European Federation for Medical Informatics*. Vol. 116. IOS Press, 2005.

[4] Imhoff, Michael, et al. *Time series analysis in intensive care medicine*. No. 1998, 01. Technical Report, SFB 475: Komplexitätsreduktion in Multivariaten Datenstrukturen, Universität Dortmund, 1998.

[5] Gersch, Will. "Spectral analysis of EEG's by autoregressive decomposition of time series." *Mathematical Biosciences* 7.1-2 (1970): 205-222.

[6] Hepworth, Joseph T., Sherry Garrett Hendrickson, and Jean Lopez. "Time series analysis of physiological response during ICU visitation." *Western Journal of Nursing Research* 16.6 (1994): 704-717.

[7] Rathlev, Niels K., et al. "Time series analysis of variables associated with daily mean emergency department length of stay." *Annals of emergency medicine* 49.3 (2007): 265-271.

[8] Jones, Spencer S., et al. "A multivariate time series approach to modeling and forecasting demand in the emergency department." *Journal of biomedical informatics* 42.1 (2009): 123-139.

[9] Kaier, K., et al. "Correlations between bed occupancy rates and Clostridium difficile infections: a time-series analysis." *Epidemiology and infection*139.03 (2011): 482-485.

[10] Gather, Ursula, Michael Imhoff, and Roland Fried. "Graphical models for multivariate time series from intensive care monitoring." *Statistics in medicine* 21.18 (2002): 2685-2701.

[11] Doric, A. G., et al. "The impact of an ICU liaison nurse service on patient outcomes." *Critical Care and Resuscitation* 10.4 (2008): 296.

[12] Kaier, K., et al. "Epidemiology meets econometrics: using time-series analysis to observe the impact of bed occupancy rates on the spread of multidrug-resistant bacteria." *Journal of Hospital Infection* 76.2 (2010): 108-113.

2) Diffusion models

Assumptions

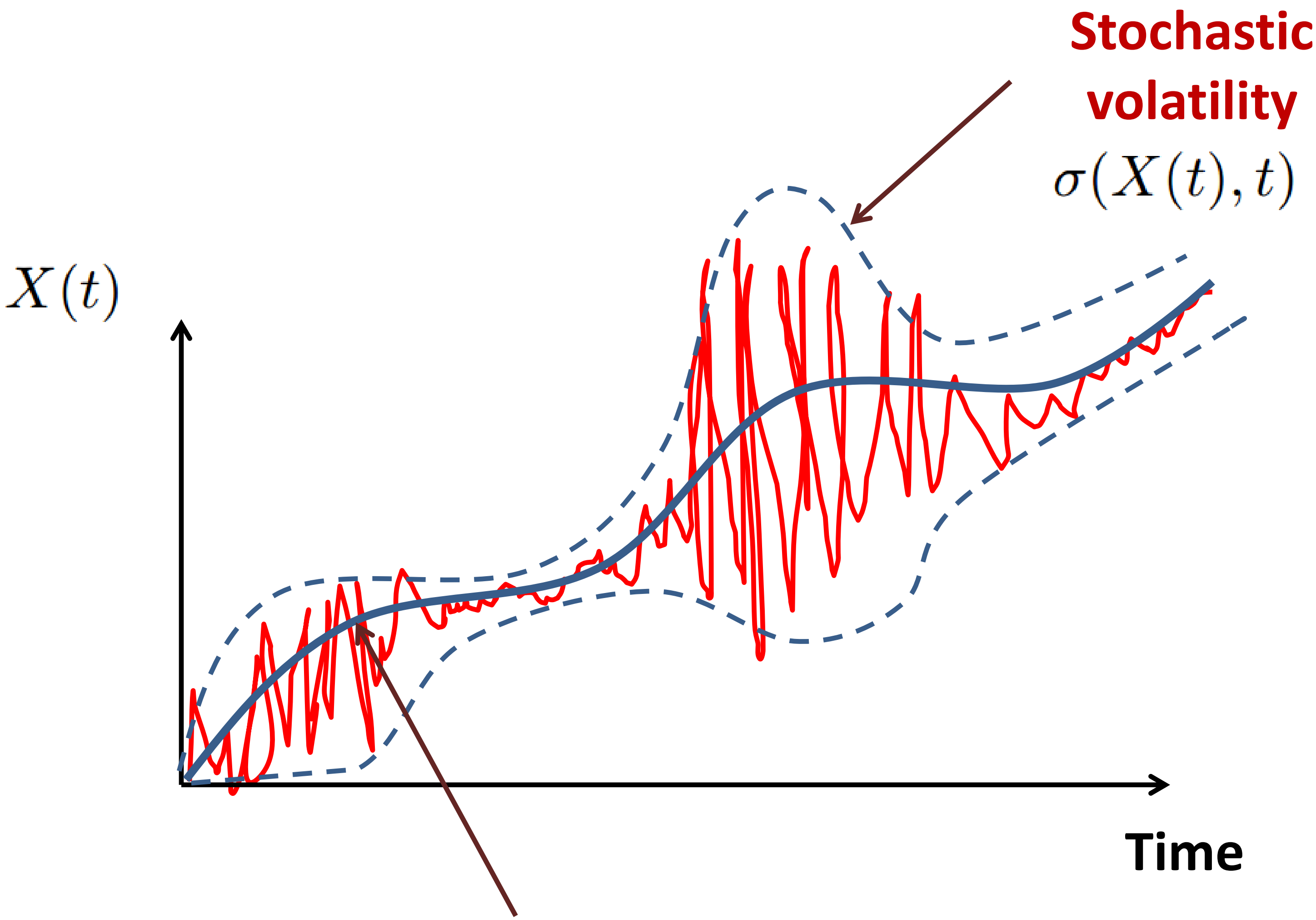
Underlying process is:

- Continuous time
- Solves a stochastic differential equation (SDE)
- Has a drift component and a diffusion component
- Usually a Markovian process

$$dX(t) = \mu(X(t), t)\, dt + \sigma(X(t), t)\, dB(t)$$

- $\mu(X(t), t)$: Drift component
- $\sigma(X(t), t)$: Diffusion component
- $B(t)$: Wiener process (standard Brownian motion)

2) Diffusion models: depiction



$\mu(X(t), t)$ **Drift component (e.g. clinical deterioration)**

2) Diffusion models

Advantages

- Models continuous time, some classical results are tractable Statistically efficient and non-stationary: only need MLE for drift and stochastic volatility

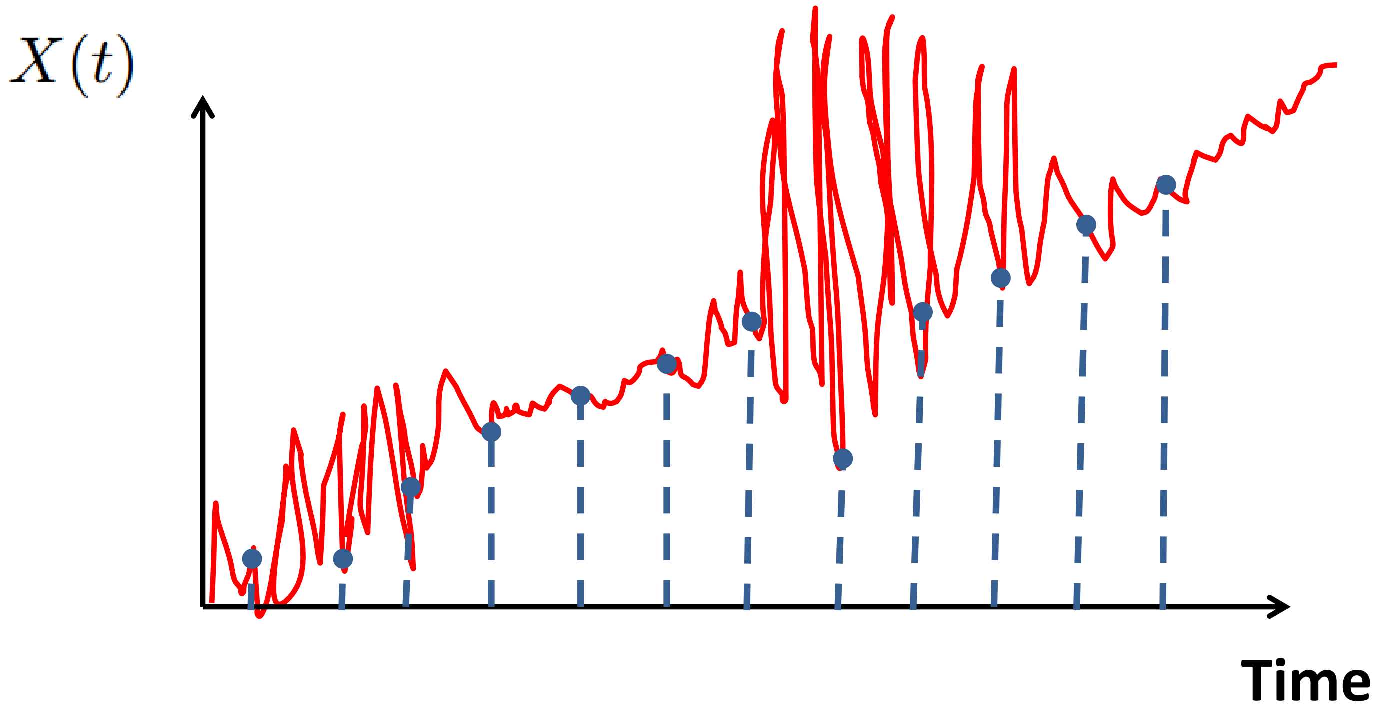
Disadvantages

- In many cases discrete time models suffice especially if sampler is exogenous. Many SDE problems are tedious and intractable.

- **Mostly limited to Markovian processes.**
- **Hard to model relational structure.**

Unexplored by ML community, intensively used in quantitative finance

2) Diffusion models: depiction



Brownian motion is the limit of many ARIMA models!

2) Diffusion models: key papers

[1] Holmes, Philip. "Optimal temporal risk assessment." (2011).

[2] Eke, Andras, et al. "Physiological time series: distinguishing fractal noises from motions." *Pflügers Archiv* 439.4 (2000): 403-415.

[3] Peng, C-K., et al. "Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series." *Chaos: An Interdisciplinary Journal of Nonlinear Science* 5.1 (1995): 82-87.

[4] Eke, Andras, et al. "Fractal characterization of complexity in temporal physiological signals." *Physiological measurement* 23.1 (2002): R1.

[5] Graversen, Svend Erik, Goran Peskir, and Albert Nikolaevich Shiryaev. "Stopping Brownian motion without anticipation as close as possible to its ultimate maximum." *Theory of Probability & Its Applications* 45.1 (2001): 41-50.

[6] Van Der Vaart, Aad, and Harry Van Zanten. "Information rates of nonparametric Gaussian process methods." *The Journal of Machine Learning Research* 12 (2011): 2095-2119.

[7] Genton, Marc G. "Classes of kernels for machine learning: a statistics perspective." *The Journal of Machine Learning Research* 2 (2002): 299-312.

[8] Jorgensen, Palle, and Feng Tian. "Discrete reproducing kernel Hilbert spaces: Sampling and distribution of Dirac-masses." *arXiv preprint arXiv:1501.02310* (2015).

[9] Heaukulani, Creighton, David Knowles, and Zoubin Ghahramani. "Beta diffusion trees." *Proceedings of the 31st International Conference on Machine Learning (ICML-14)*. 2014.

[11] Zhang, Chao, and Dacheng Tao. "Risk bounds of learning processes for lévy processes." *The Journal of Machine Learning Research* 14.1 (2013): 351-376.

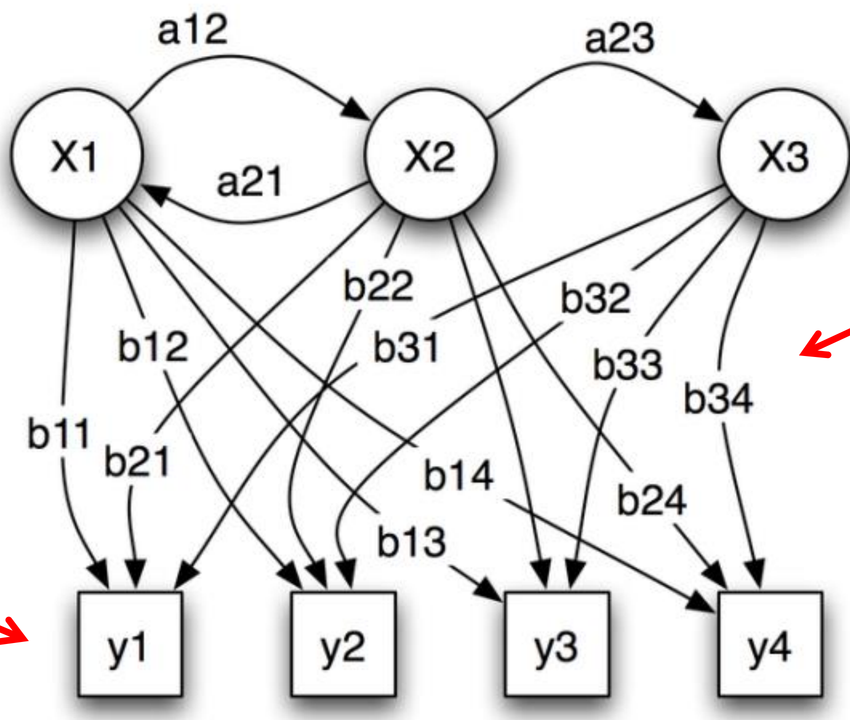
[12] Li, Yuxi, Csaba Szepesvari, and Dale Schuurmans. "Learning exercise policies for american options." *International Conference on Artificial Intelligence and Statistics*. 2009.

3) Hidden Markov Models (HMMs)

Assumptions

States are not observable, and what is observed is a set of outputs that are observed with a different probability given each state.

States (unobservable)

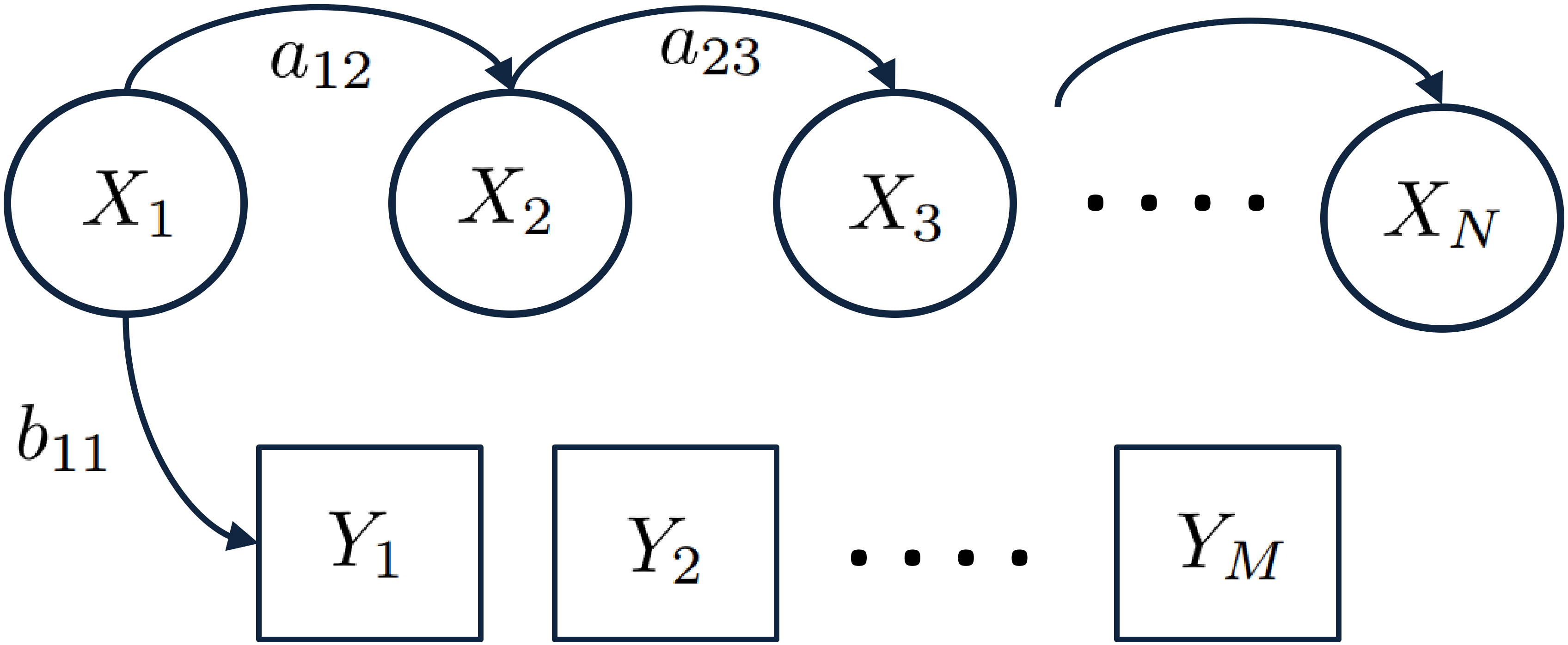


Output-state distribution

Outputs (observable)

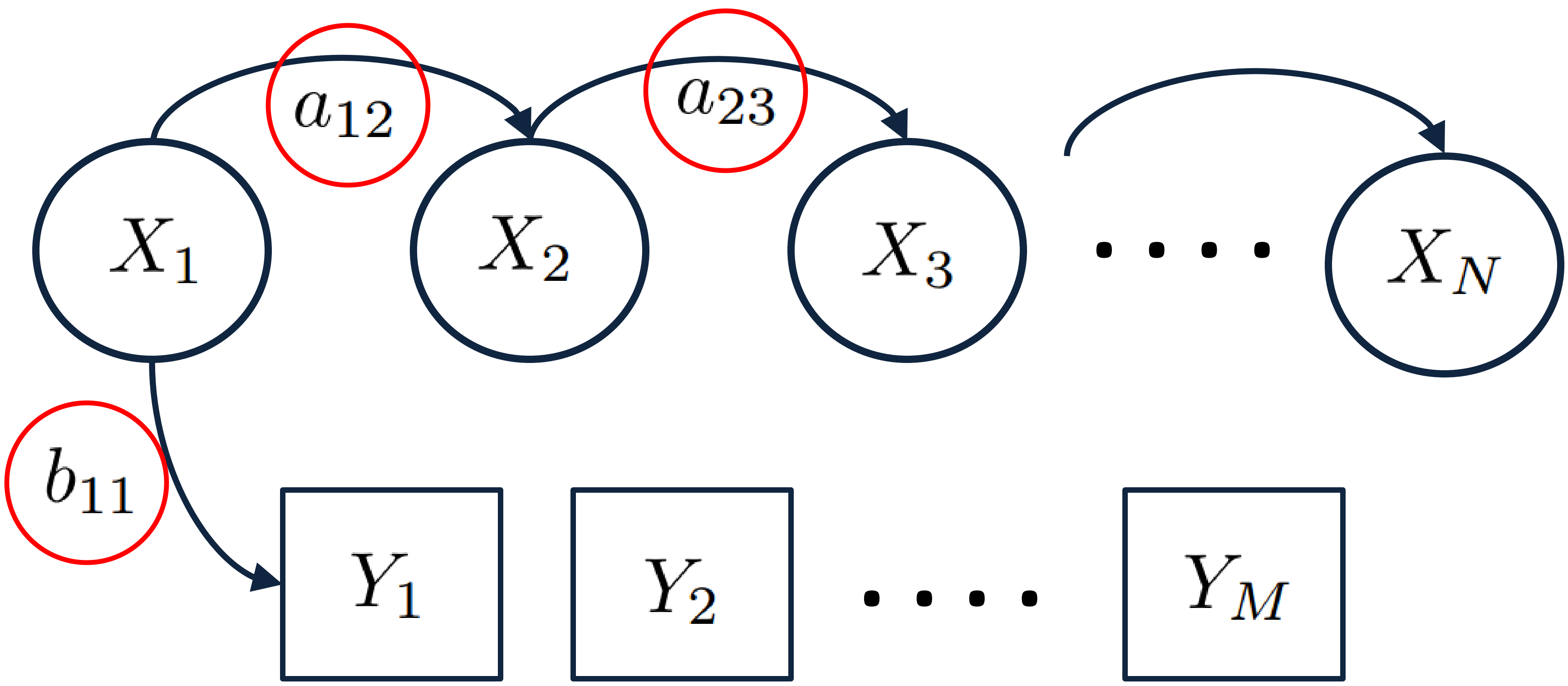
3) HMM simulation

- **Observation is a probabilistic function of the state**
- **HMM is a doubly embedded stochastic process**



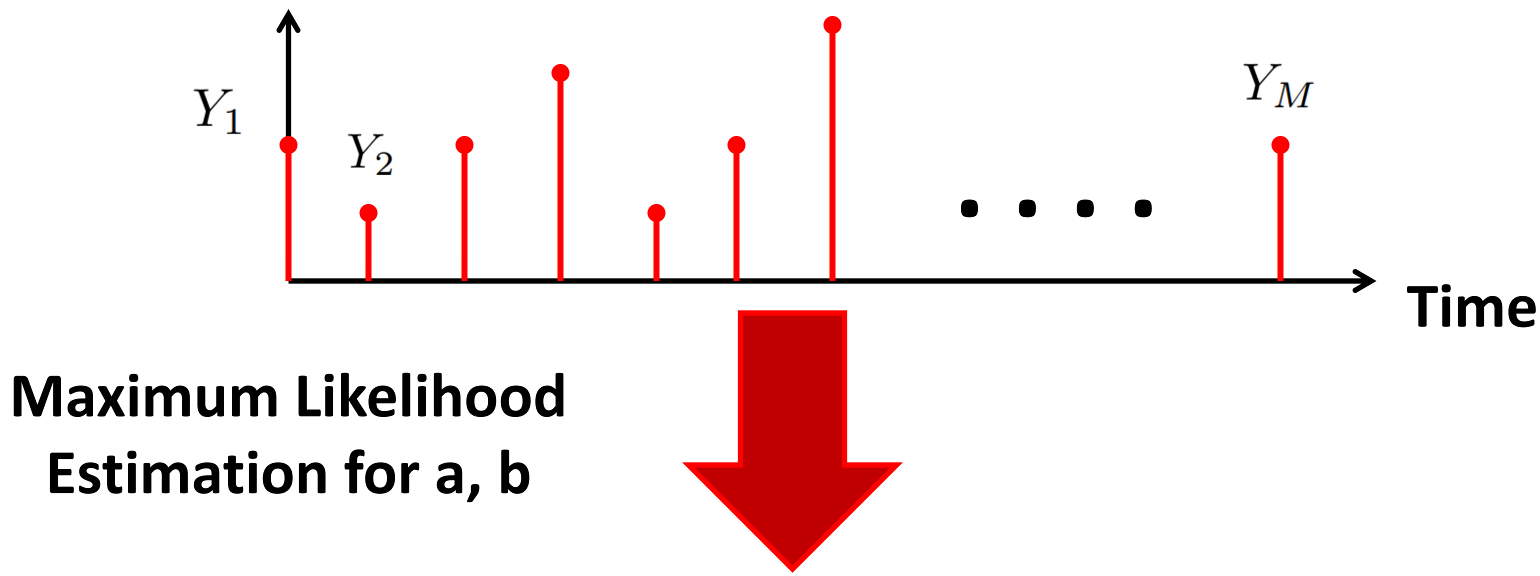
3) HMM learning

- **Learning an HMM is equivalent to learning the state-transition probabilities (a_{ij}) and the emission probabilities (b_{ik})**



3) HMM learning

- **Learning an HMM is equivalent to learning the state-transition probabilities (a_{ij}) and the emission probabilities (b_{ik}) given an output sequence.**



$$\arg\max_{a,b} \mathbb{P}\left(Y_1, Y_2, Y_3, \ldots, Y_M \,\middle|\, \{a_{ij}\}, \{b_{ij}\}\right)$$

3) HMM learning using dynamic programming

Algorithms for learning HMMs

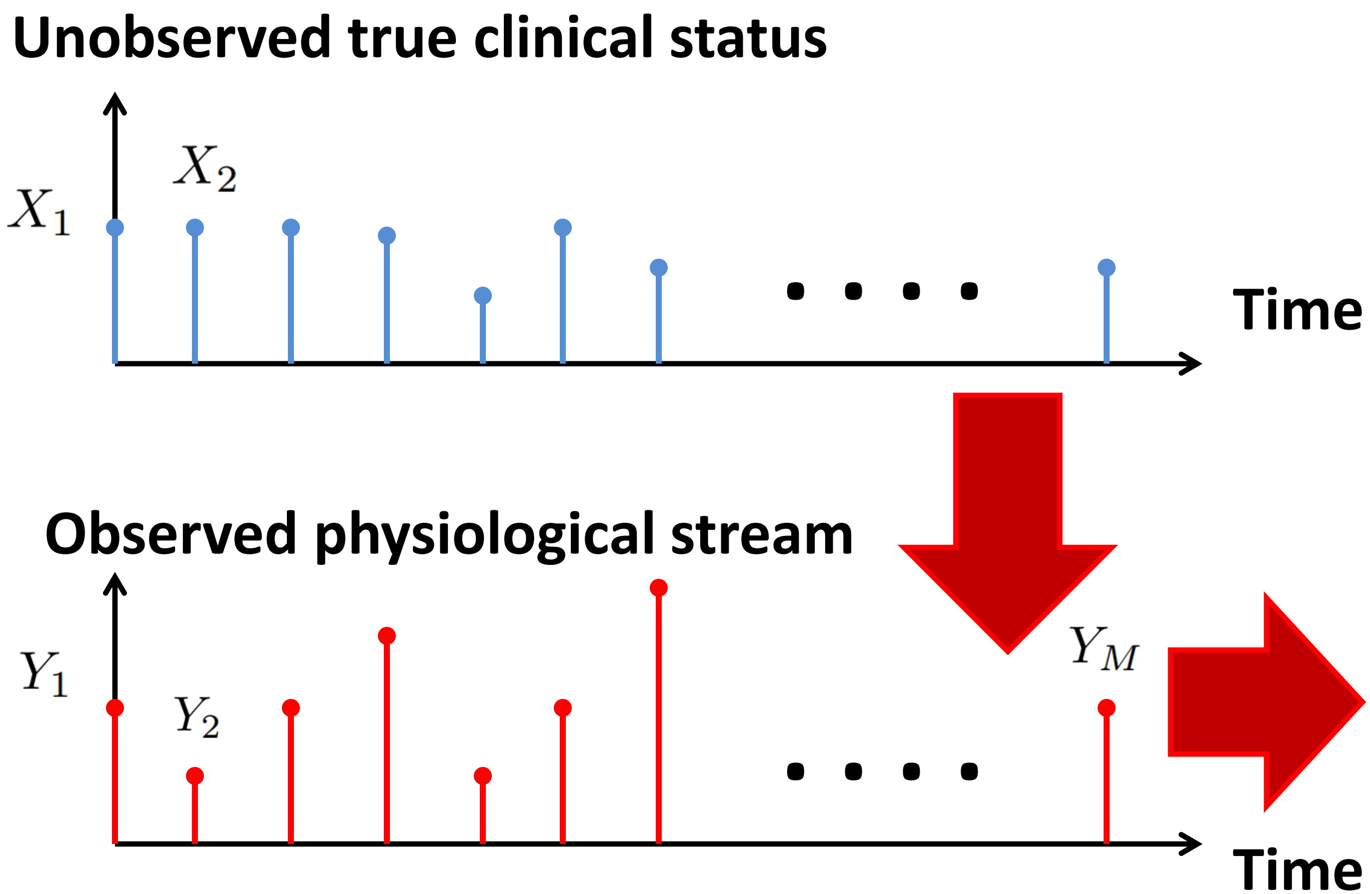
1- Brute force: enumerate all output sequences and compute their likelihood -> exponential complexity!

2- Viterbi algorithm (dynamic programming): same idea as Viterbi decoders in convolutional codes.

Used if there is a known state-space that is not observed, and that has a non-injective map to an observed output sequence.

3) HMM learning using dynamic programming

HMMs in a medical context (e.g. Martin's problem)



Learn model by Viterbi, Decoder emits a risk score!

3) HMM control

Actions + rewards

Markov Chain → MDP

Actions + rewards

HMM → POMDP

3) HMMs: key papers

[1] Zhang, Yongyue, Michael Brady, and Stephen Smith. "Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm." *Medical Imaging, IEEE Transactions on* 20.1 (2001): 45-57.

[2] Andreão, Rodrigo Varejão, Bernadette Dorizzi, and Jérôme Boudy. "ECG signal analysis through hidden Markov models." *Biomedical Engineering, IEEE Transactions on* 53.8 (2006): 1541-1549.

[3] Coast, Douglas A., et al. "An approach to cardiac arrhythmia analysis using hidden Markov models." *Biomedical Engineering, IEEE Transactions on* 37.9 (1990): 826-836.

[4] Oates, Tim, Laura Firoiu, and Paul R. Cohen. "Using dynamic time warping to bootstrap HMM-based clustering of time series." *Sequence Learning*. Springer Berlin Heidelberg, 2000. 35-52.

[5] Oates, Tim, Laura Firoiu, and Paul R. Cohen. "Clustering time series with hidden markov models and dynamic time warping." *Proceedings of the IJCAI-99 workshop on neural, symbolic and reinforcement learning methods for sequence learning*. Sweden Stockholm, 1999.

[6] Oates, Tim, Laura Firoiu, and Paul R. Cohen. "Clustering time series with hidden markov models and dynamic time warping." *Proceedings of the IJCAI-99 workshop on neural, symbolic and reinforcement learning methods for sequence learning*. Sweden Stockholm, 1999.

[7] Pavlovic, Vladimir, Brendan J. Frey, and Thomas S. Huang. "Time-series classification using mixed-state dynamic Bayesian networks." *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.*. Vol. 2. IEEE, 1999.

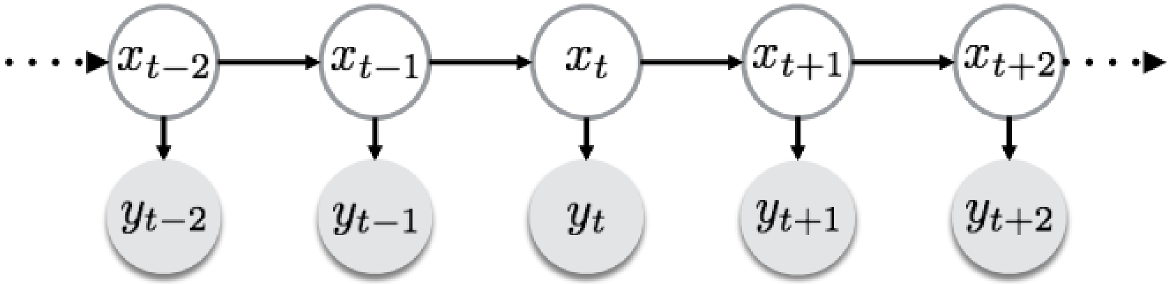
[8] Li, Sheng-Tun, and Yi-Chung Cheng. "A stochastic HMM-based forecasting model for fuzzy time series." *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on* 40.5 (2010): 1255-1266.

[9] Listgarten, Jennifer, et al. "Multiple alignment of continuous time series."*Advances in neural information processing systems*. 2004.

4) State-space model

Assumptions

- Continuous state-space
- This leads to linearity assumptions for mapping state transitions and states to outputs.

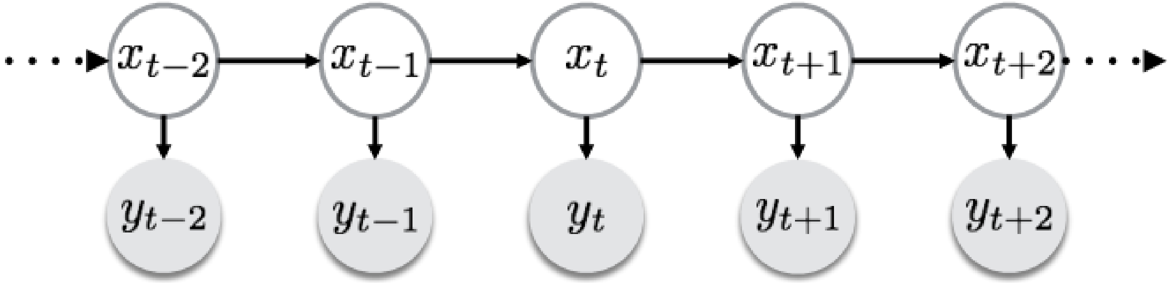


$$x_t = Ax_{t-1} + e_t \quad e_t \sim N(0, \Sigma)$$

$$y_t = Cx_t + w_t \quad w_t \sim N(0, R)$$

4) State-space model

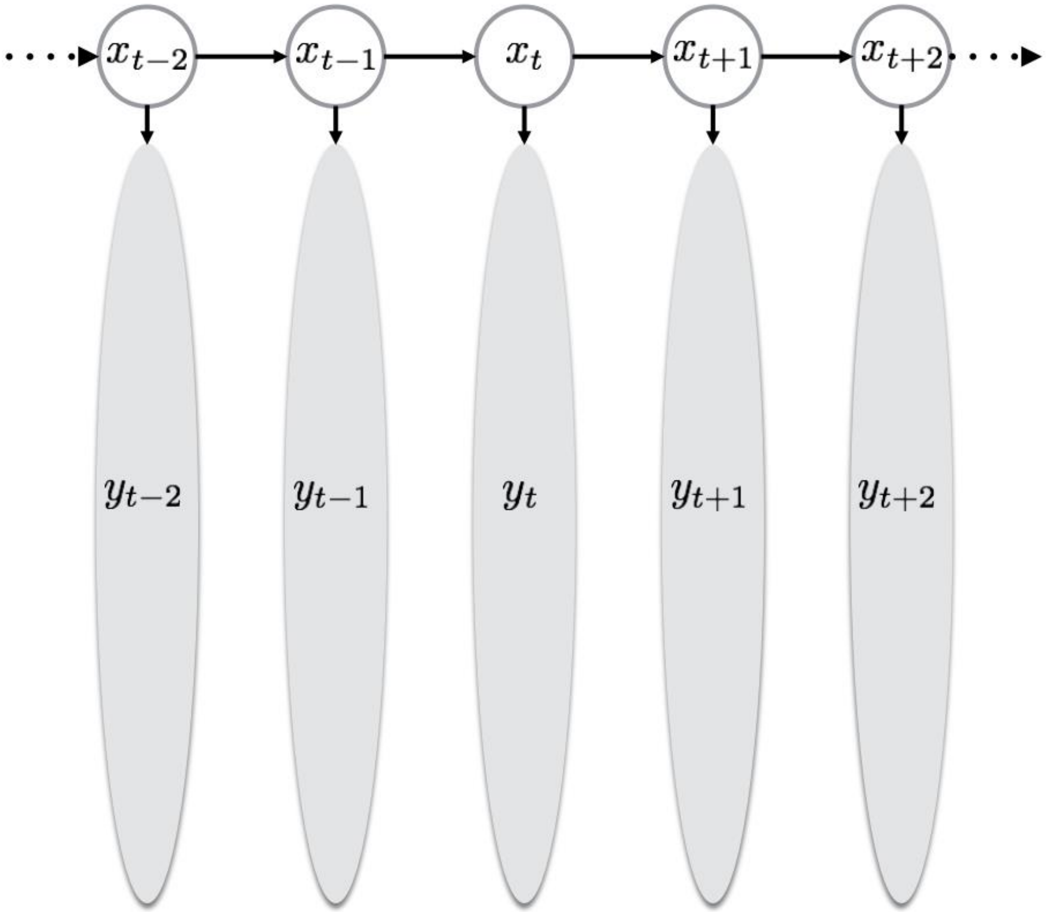
Learning task: finding the matrices mapping states to outputs, and transition matrices across states + covariance of Gaussian noise



$$x_t = Ax_{t-1} + e_t \quad e_t \sim N(0, \Sigma)$$
$$y_t = Cx_t + w_t \quad w_t \sim N(0, R)$$

4) State-space model

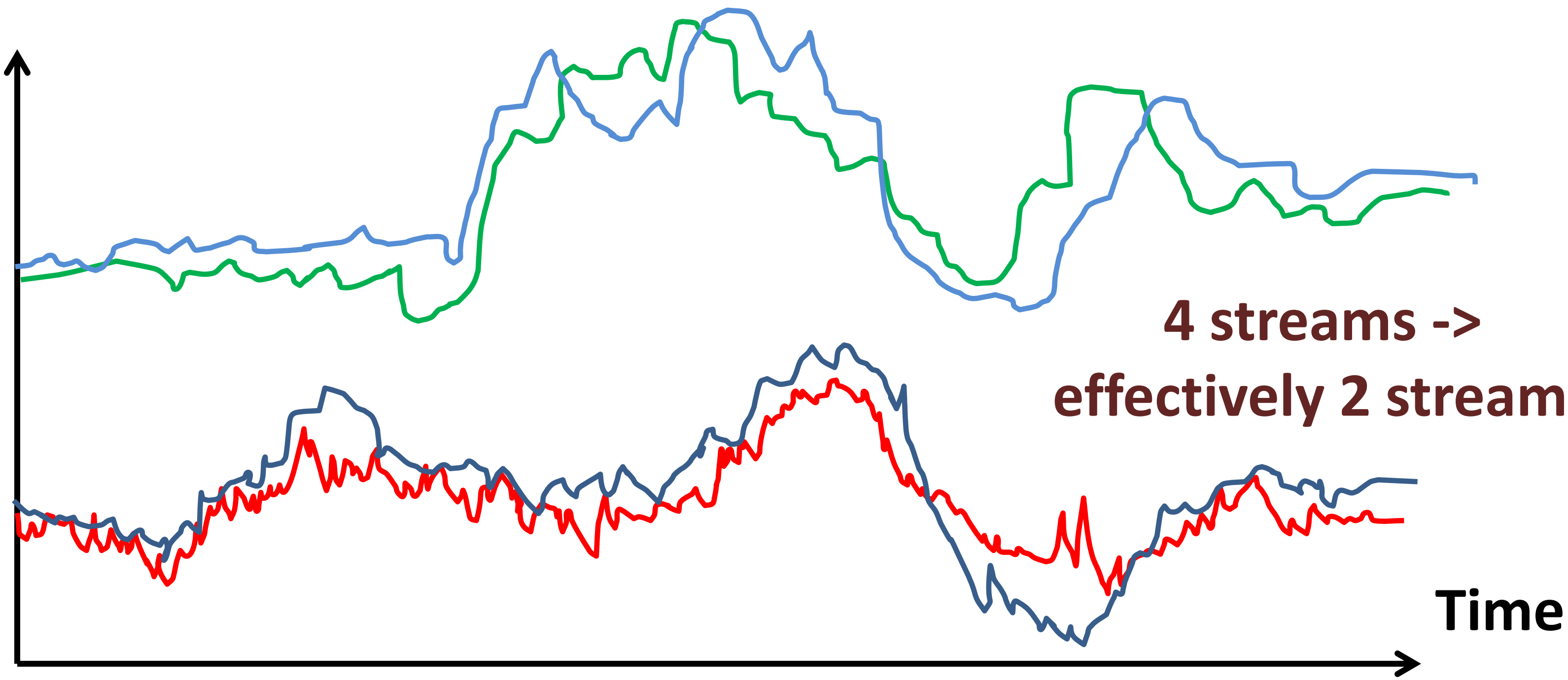
High-dimensional time series: Large matrices A and C



4) State-space model

Some temporal streams are redundant:

How to reduce the dimensionality and capture dynamics?
Redundancy can be time-varying!!



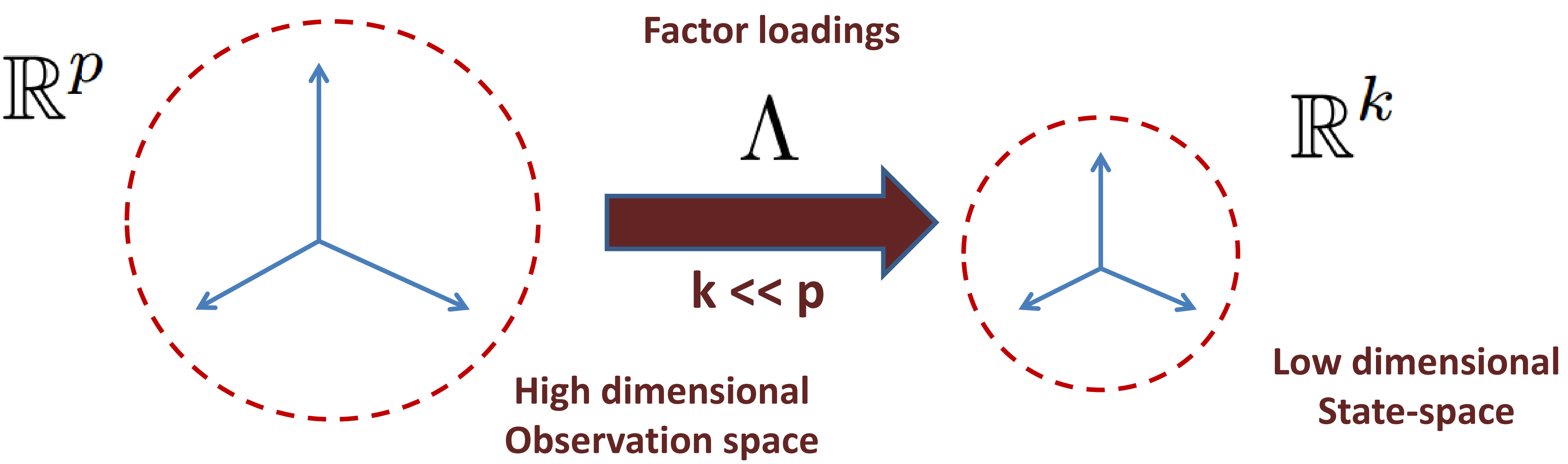
4 streams -> effectively 2 streams

5) Dynamic Latent Factor Models

Goal: Embed high dimensional time series into a lower dimensional space and maintain the dynamics

Dynamic functional connectivity (neuroscience) -> Time varying correlations

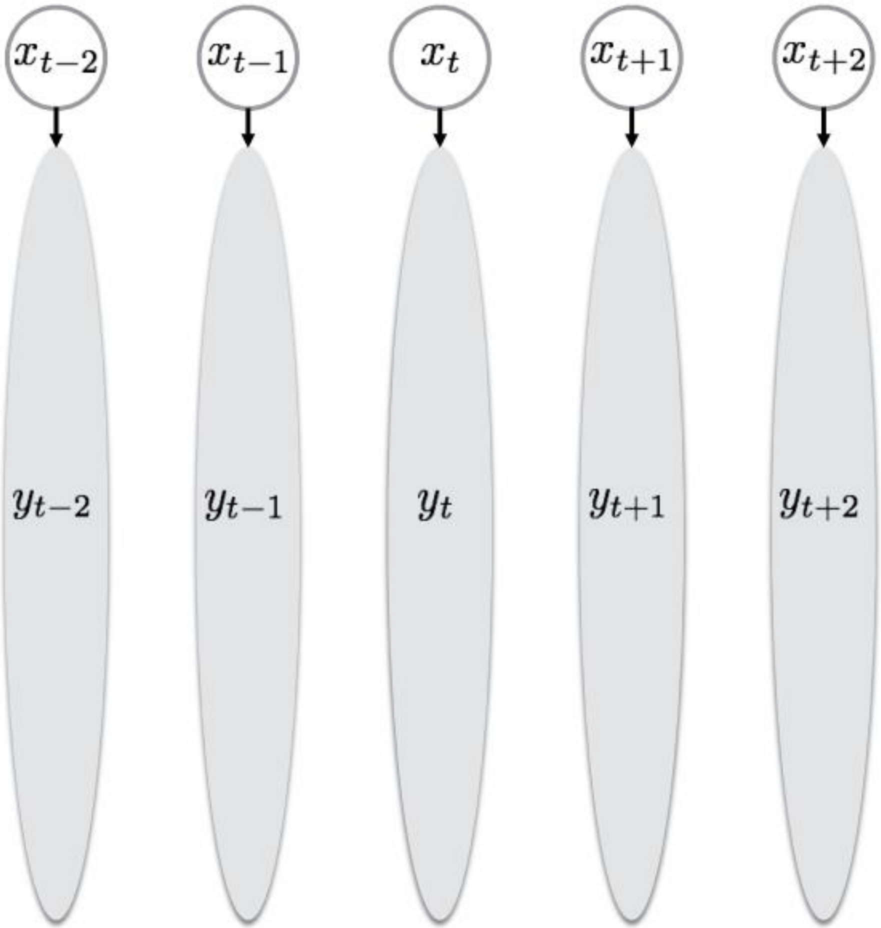
Models that seek low dimensional embedding of dynamics



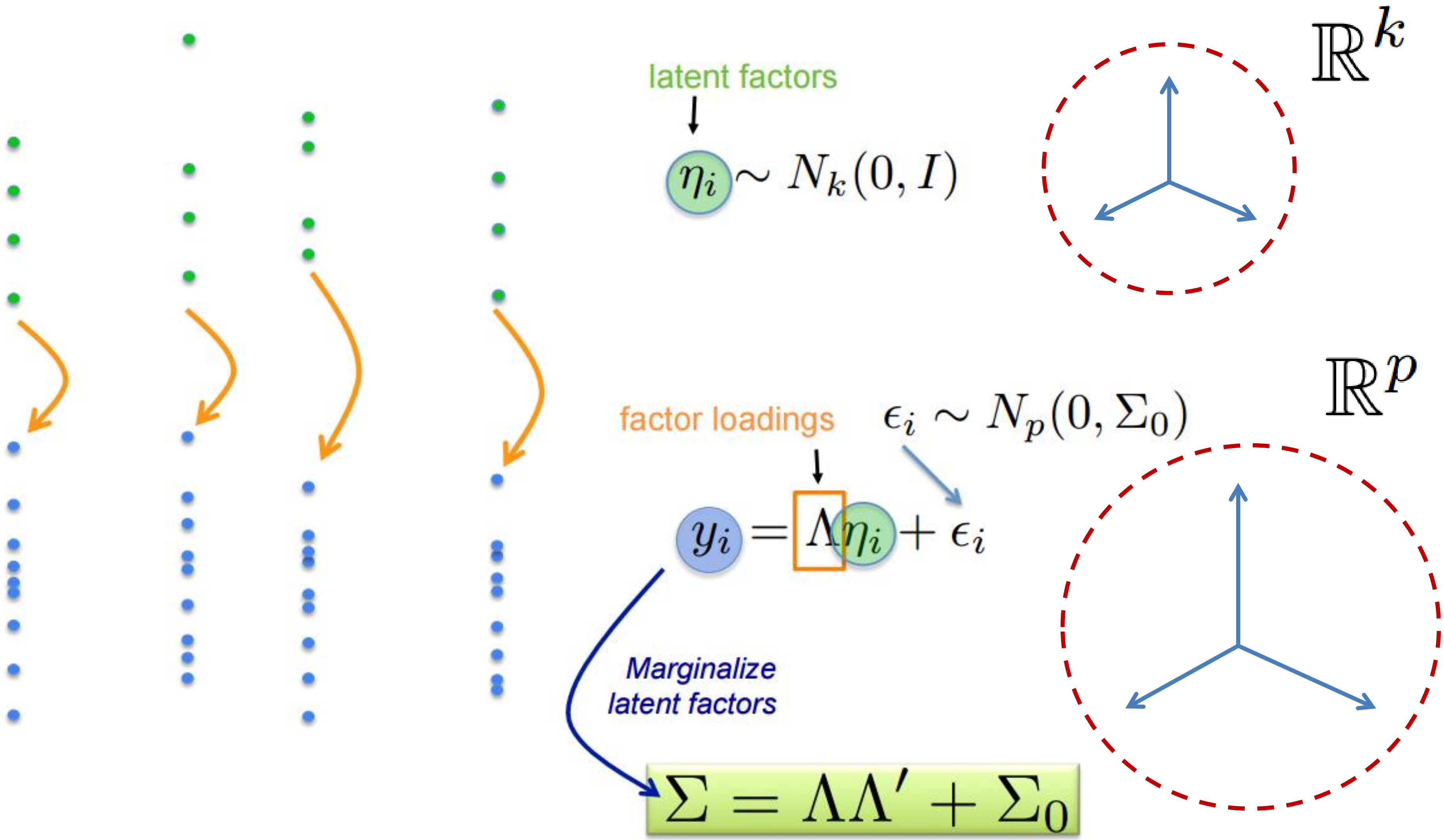
5) Dynamic Latent Factor Models

i.i.d Sequence with no dynamics =>

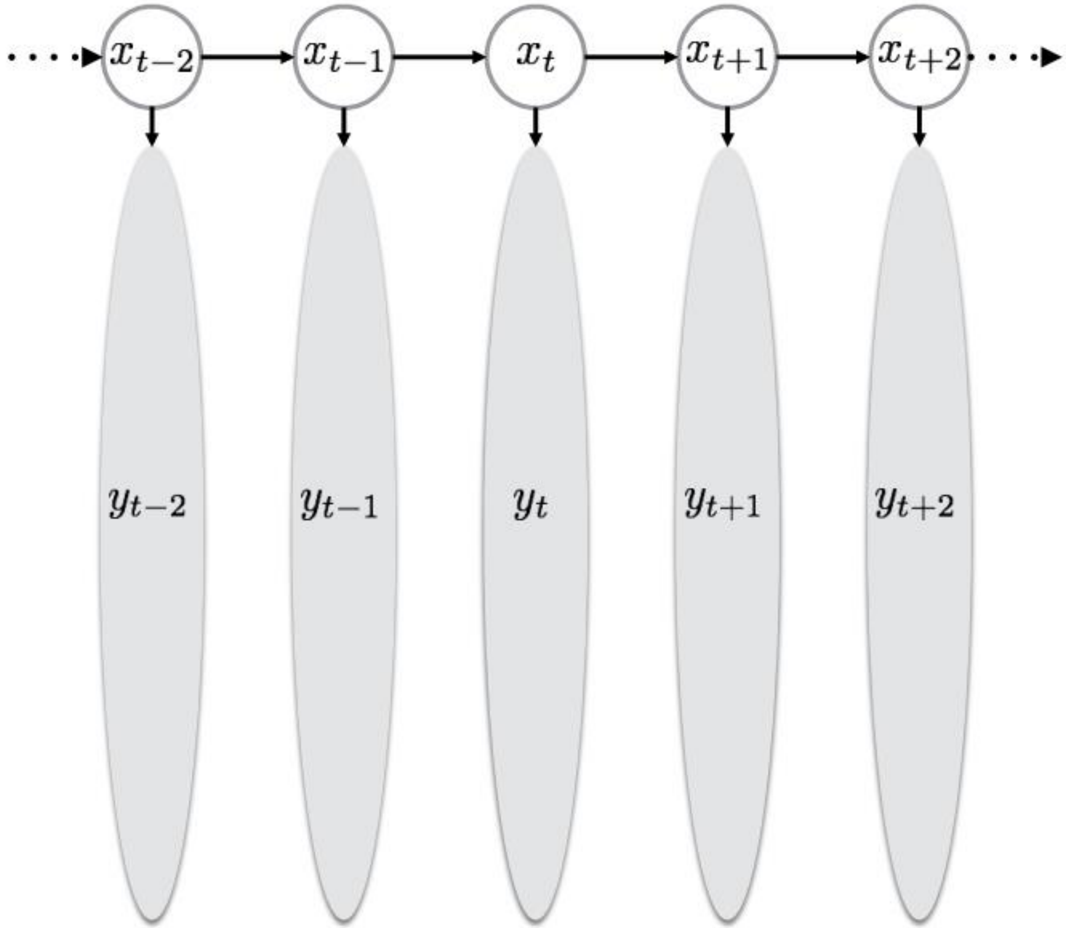
Latent factor model



5) Dynamic Latent Factor Models



5) Dynamic Latent Factor Models

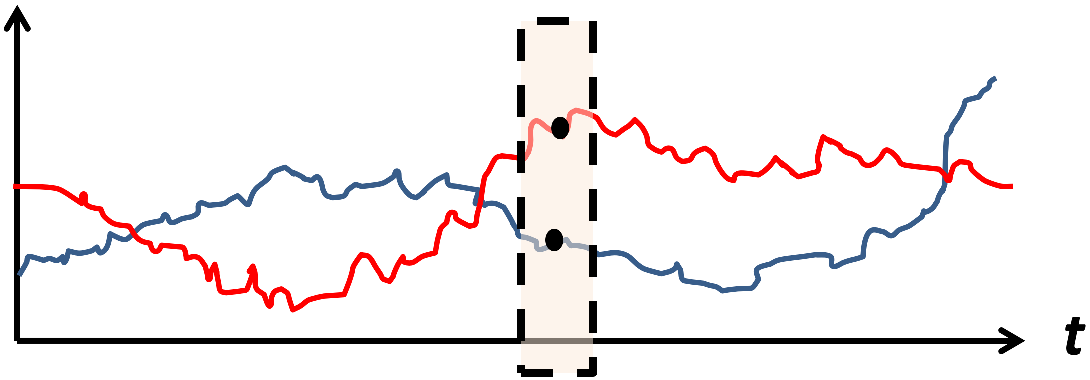


Incorporating dynamics =>

Dynamic latent factor model

5) Dynamic Latent Factor Models

2-dimensional state-space dynamics

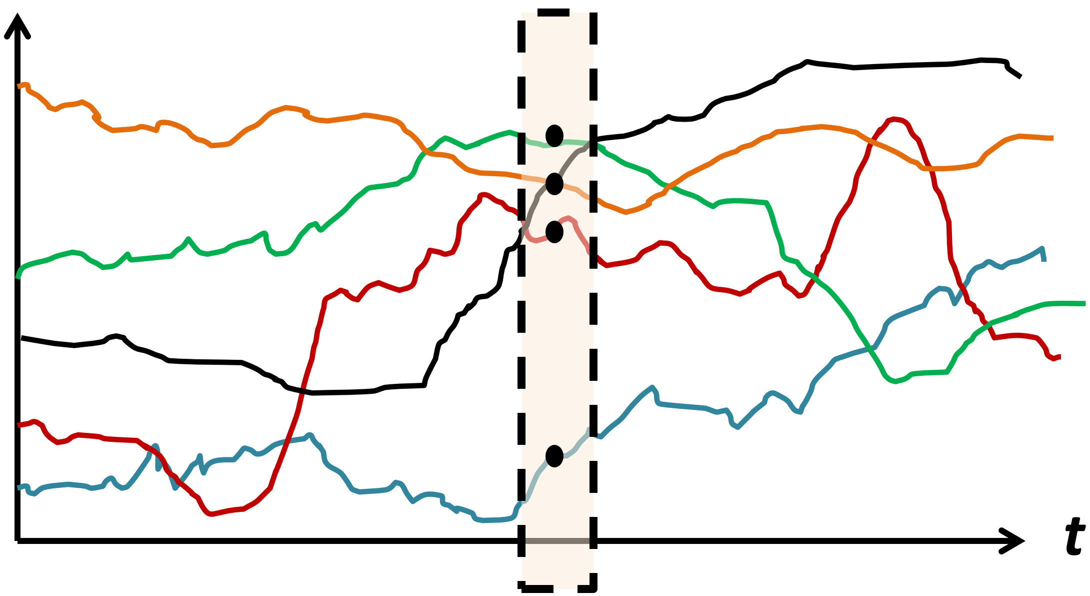


$$\eta_t = \Phi \eta_{t-1} + \nu_t$$

$\nu_t \sim N_k(0, I)$

Evolution of latent factors ($\Phi \eta_{t-1}$)

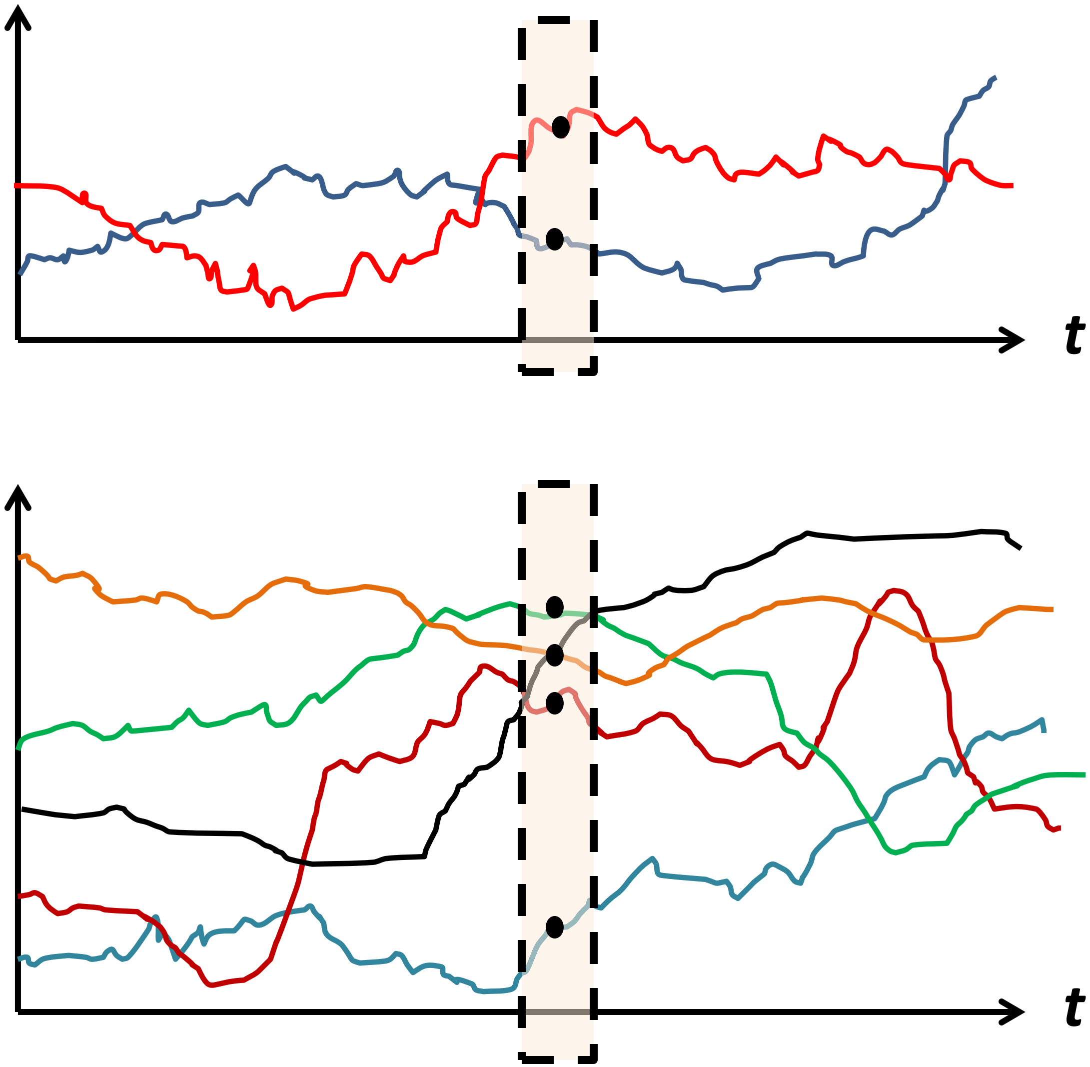
5-dimensional observation-space dynamics



$$y_t = \Lambda \eta_t + \epsilon_t$$

$\epsilon_t \sim N_p(0, \Sigma_0)$

5) Dynamic Latent Factor Models



Correlation pattern is still Fixed over time!

$$y_t \sim N(0, \Sigma)$$

$$\Sigma = \Lambda \Sigma_\eta \Lambda' + \Sigma_0$$

(Σ_η: $\Gamma_\eta(0)$)

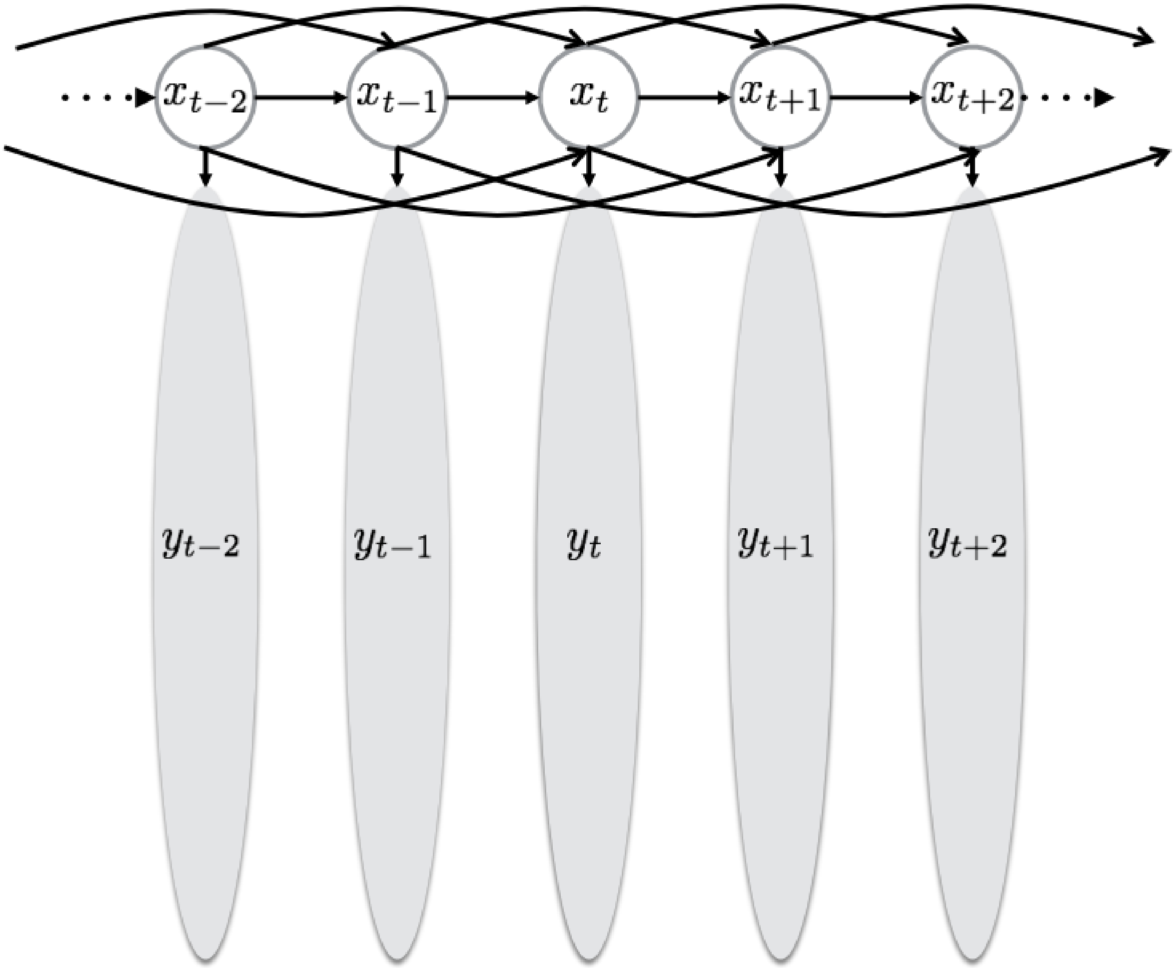
Lag covariance = WSS

$$\Gamma_y(h) = \text{cov}(y_t, y_{t+h})$$
$$= \Lambda \Gamma_\eta(h) \Lambda' \quad h > 0$$

5) Dynamic Latent Factor Models

Complex dynamics=>

Nonparametric latent factor evolution



5) Dynamic Latent Factor Models

Complex dynamics=>

Nonparametric latent factor evolution

Usually a Gaussian process is used for nonparametric evolution

Time index

$N_k(0, I)$

$$\eta_t = \psi(x_t) + \nu_t$$

Nonparametric evolution of latent factors

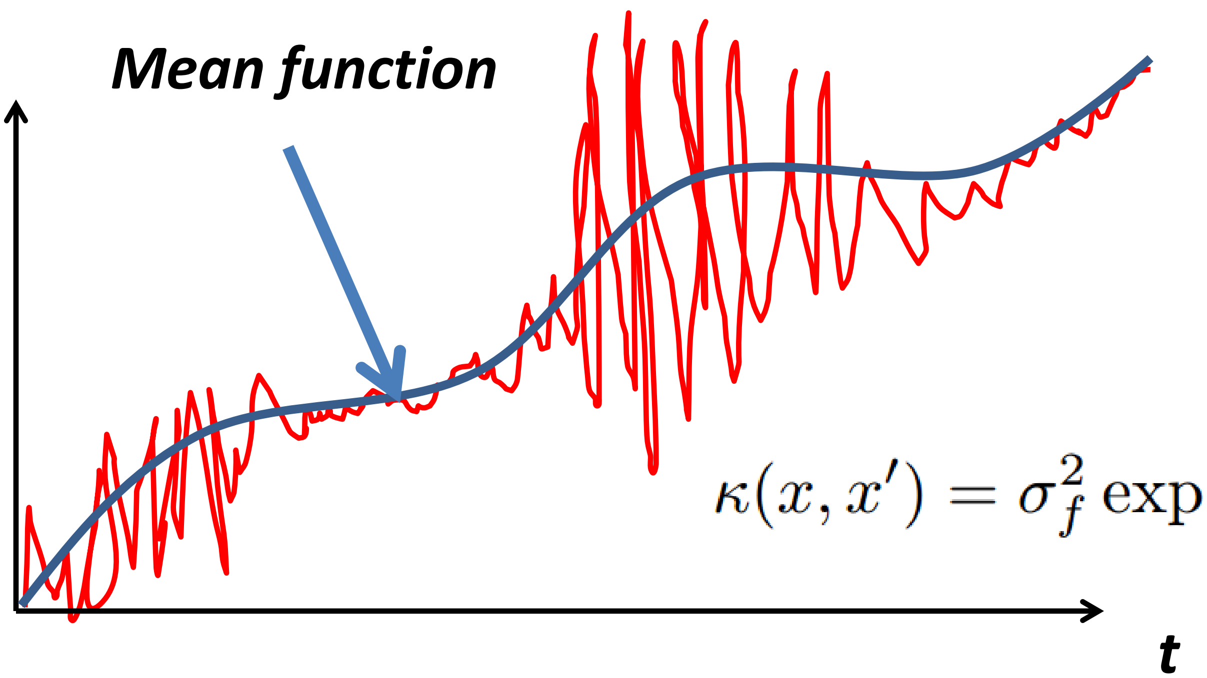
$$y_t = \Lambda \eta_t + \epsilon_t$$

$N_p(0, \Sigma_0)$

What is a Gaussian Process?

- **Distribution over functions f ~ GP (m, K)**
- **m: mean function, K: covariance Kernel**

Brownian motion is a GP with Matern Kernel



$$\kappa(x, x') = \sigma_f^2 \exp\left(-\frac{1}{2\ell^2}(x - x')^2\right)$$

5) Dynamic Latent Factor Models

Fit the latent factor process using GP regression => Get f(.)

$$N_k(0, I)$$

$$\eta_t = f(\eta_{1:t-1}) + \nu_t$$

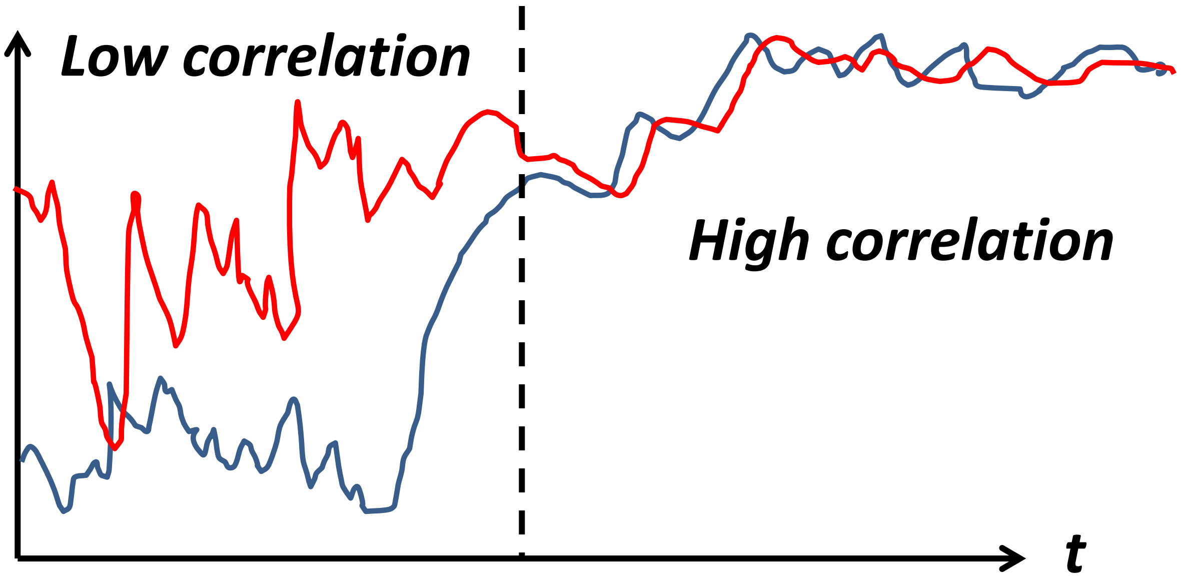
Evolution of latent factors

$$y_t = \Lambda \eta_t + \epsilon_t$$

$$N_p(0, \Sigma_0)$$

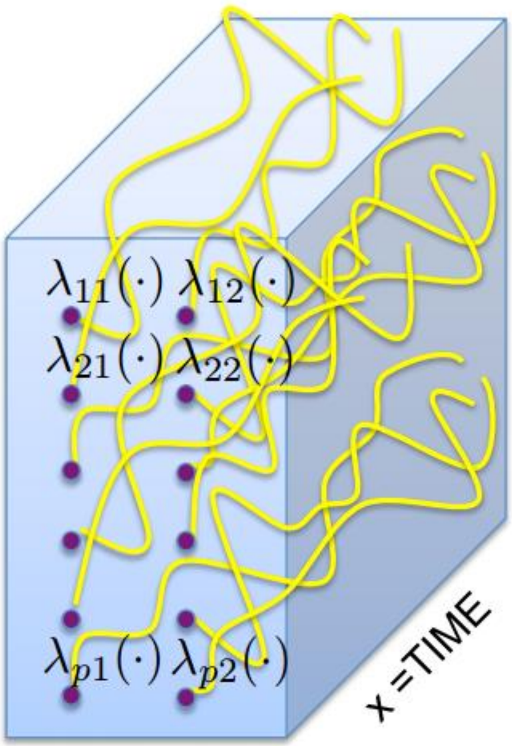
5) Dynamic Latent Factor Models

We modeled EVOLUTION by a GP, but How to capture the changing correlation pattern (RELATIONAL STRUCTURE)?



5) Dynamic Latent Factor Models

Model every entry in Latent factor matrix by a GP



$p \times k$ array of processes over time

$k << p$

$p \times k$

$$\Sigma(x) = \Lambda(x)\Lambda(x)' + \Sigma_0$$

5) Dynamic Latent Factor Models

GP to model evolution and dynamics

$N_k(0, I)$

$$\eta_t = f(\eta_{1:t-1}) + \nu_t$$

Evolution of latent factors

GP to model latent factor process

$$y_t = \Lambda \eta_t + \epsilon_t$$

$N_p(0, \Sigma_0)$

5) Dynamic Latent Factor Models

Conclusion: Dynamic Latent Factor Model can capture complicated dynamics and encapsulates most of other models

5) Dynamic Latent Factor Model: key papers

[1] Zhang, Yongyue, Michael Brady, and Stephen Smith. "Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm." *Medical Imaging, IEEE Transactions on* 20.1 (2001): 45-57.

[2] Andreão, Rodrigo Varejão, Bernadette Dorizzi, and Jérôme Boudy. "ECG signal analysis through hidden Markov models." *Biomedical Engineering, IEEE Transactions on* 53.8 (2006): 1541-1549.

[3] Coast, Douglas A., et al. "An approach to cardiac arrhythmia analysis using hidden Markov models." *Biomedical Engineering, IEEE Transactions on* 37.9 (1990): 826-836.

[4] Oates, Tim, Laura Firoiu, and Paul R. Cohen. "Using dynamic time warping to bootstrap HMM-based clustering of time series." *Sequence Learning*. Springer Berlin Heidelberg, 2000. 35-52.

[5] Oates, Tim, Laura Firoiu, and Paul R. Cohen. "Clustering time series with hidden markov models and dynamic time warping." *Proceedings of the IJCAI-99 workshop on neural, symbolic and reinforcement learning methods for sequence learning*. Sweden Stockholm, 1999.

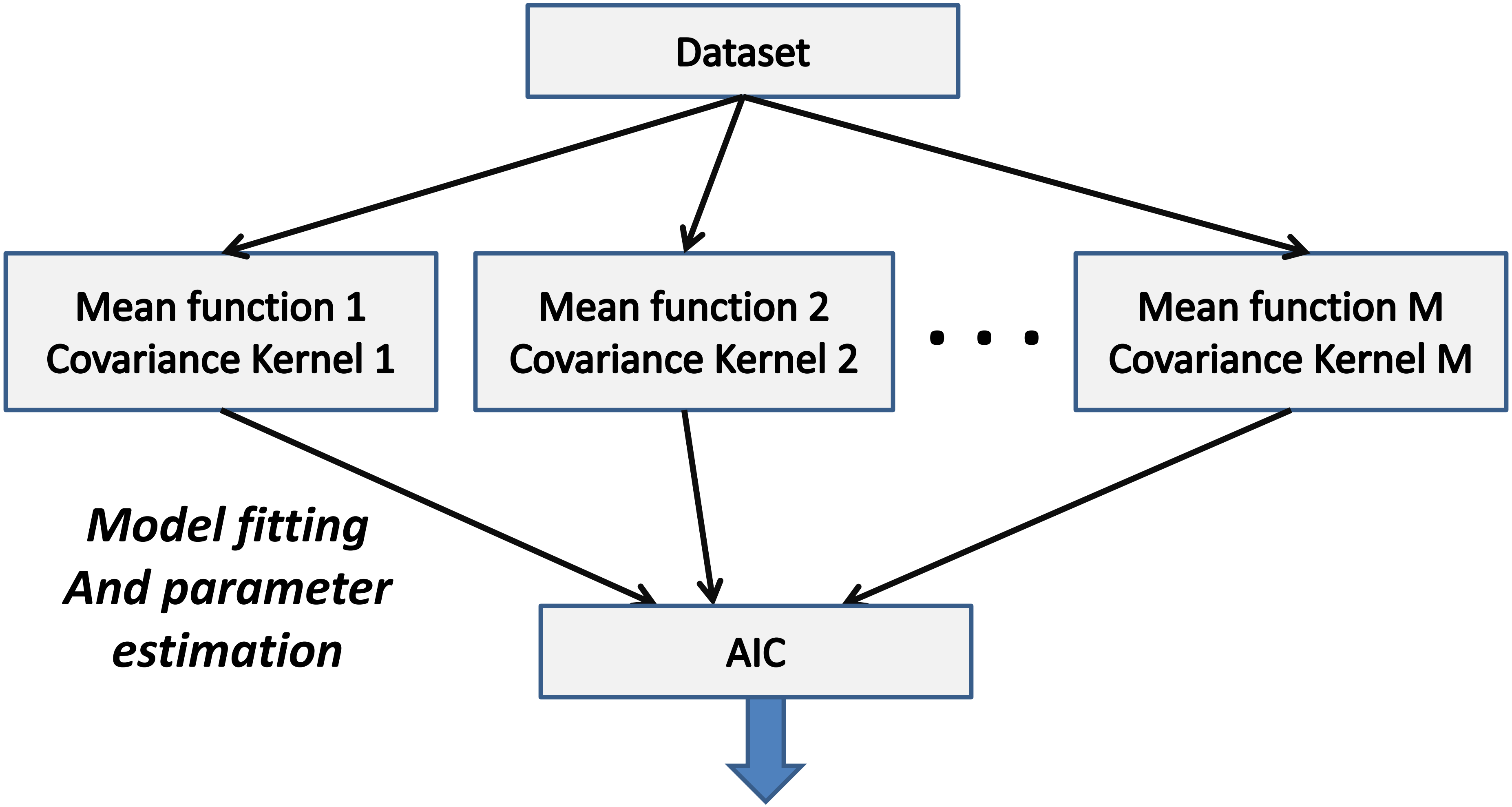
[6] Oates, Tim, Laura Firoiu, and Paul R. Cohen. "Clustering time series with hidden markov models and dynamic time warping." *Proceedings of the IJCAI-99 workshop on neural, symbolic and reinforcement learning methods for sequence learning*. Sweden Stockholm, 1999.

[7] Pavlovic, Vladimir, Brendan J. Frey, and Thomas S. Huang. "Time-series classification using mixed-state dynamic Bayesian networks." *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.*. Vol. 2. IEEE, 1999.

[8] Li, Sheng-Tun, and Yi-Chung Cheng. "A stochastic HMM-based forecasting model for fuzzy time series." *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on* 40.5 (2010): 1255-1266.

[9] Listgarten, Jennifer, et al. "Multiple alignment of continuous time series."*Advances in neural information processing systems*. 2004.

Model Selection



Select best model with a desired model confidence interval

Further topics to discuss…..

- **Dynamic Bayesian networks:** modeling cause and effect
- **Real-time Cox regression:** survival analysis based on time series
- **Structure learning:** learning graphs of patients, treatments, etc, based on time series
- **Time series clustering:** personalization!