Guided Tour of Machine Learning in Finance

Week 4: Reinforcement Learning

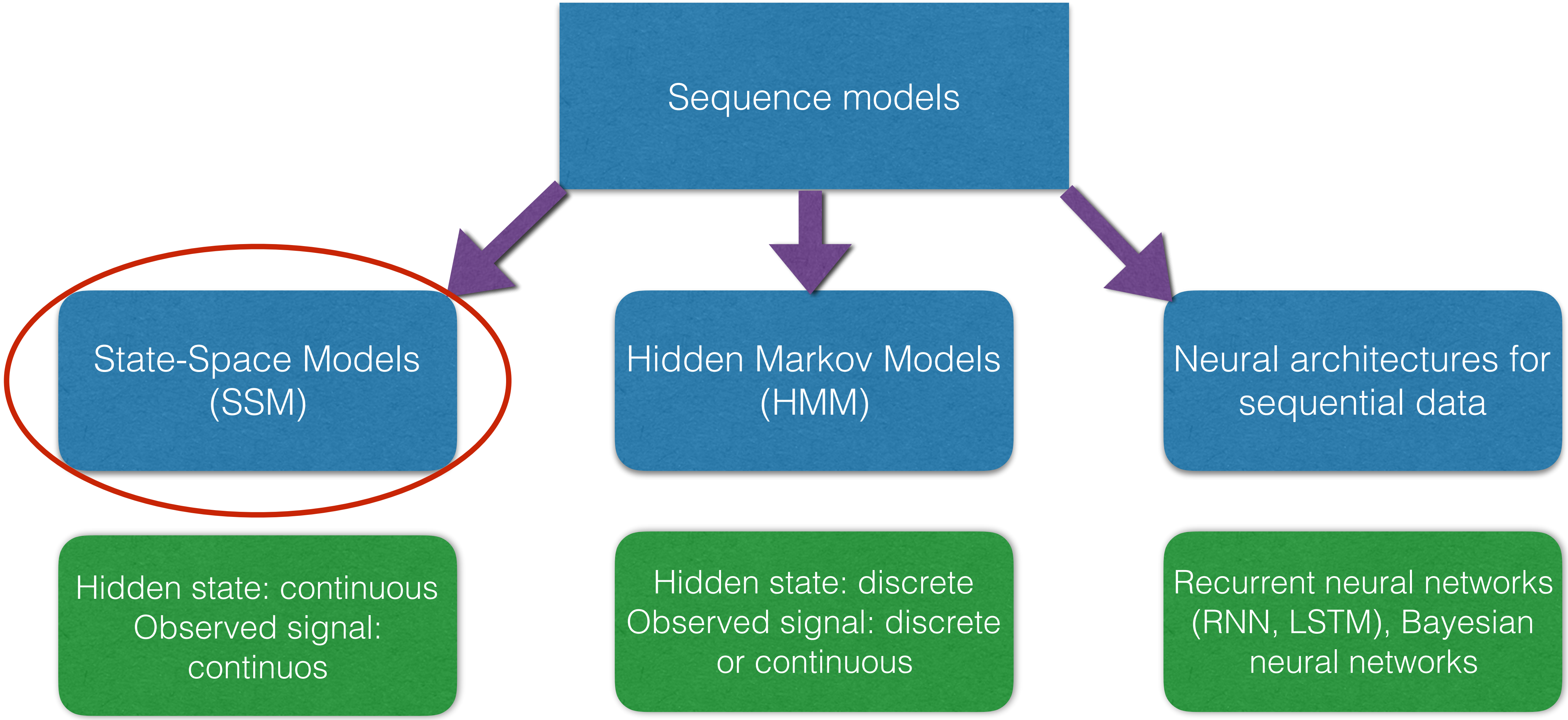
4-1-3-State-Space Models

Igor Halperin

NYU Tandon School of Engineering, 2017

Sequence models

Parametric (SSM, HMM) vs non-parametric (neural) of sequence modeling with a hidden state $p(y \mid \mathbf{x})$



State-Space Models (SSM)

The last: build a probabilistic model for the observable signal $y = \left\{ y^{(t)} \right\}_{t=1}^{T}$

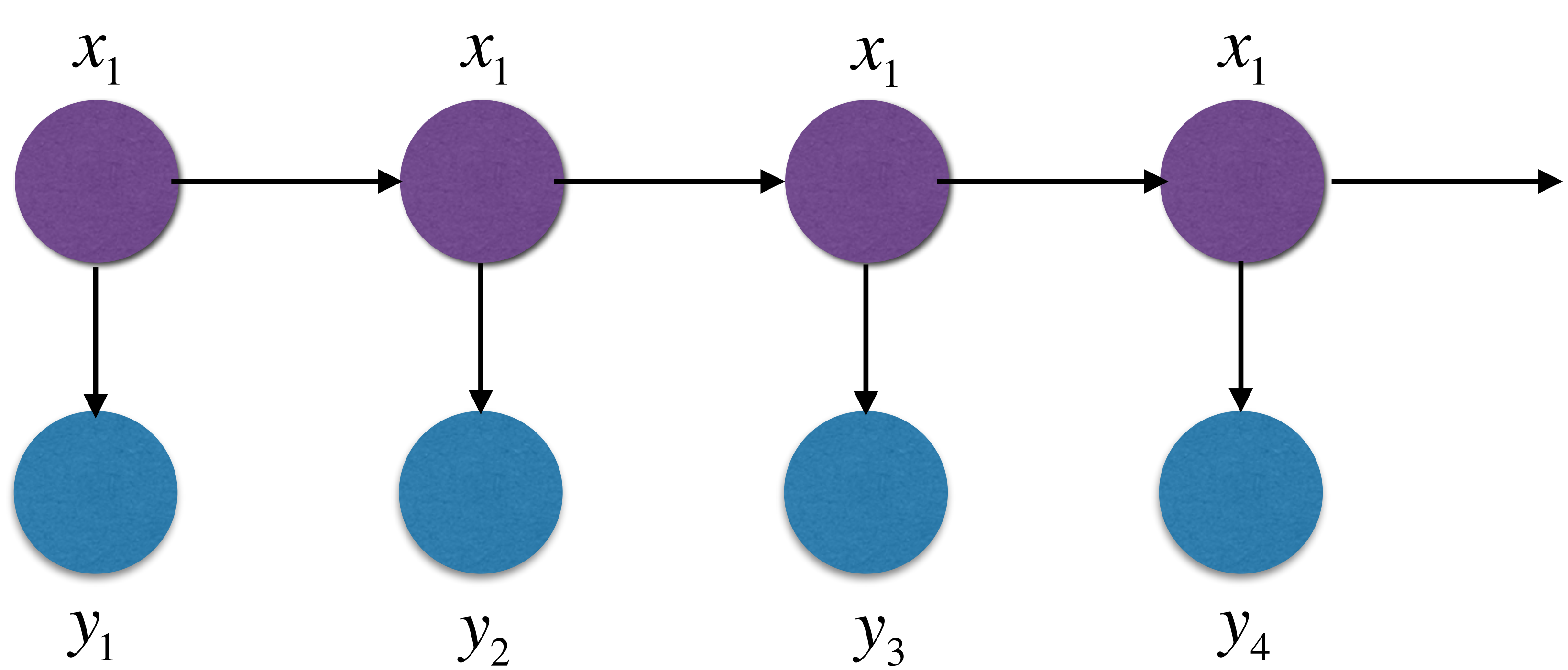
SSM = dynamic latent variables models with an **continuous** hidden state $x^{(t)}$

- The hidden state $x^{(t)}$ captures the dynamics of the system, filters noise out
- Used as a conditioning variable for predictions $p\left(y^{(t)} \mid y^{(t-1)}\right) \rightarrow p\left(y^{(t)} \mid x^{(t)}\right)$

The dynamics is first-order Markov in the hidden state:

$$p(x_{1:T}, y_{1:T} \mid \theta) = \prod_{t=1}^{T} p(x_t \mid x_{t-1}, \theta)\, p(y_t \mid x_t, \theta)$$

- Hidden states x have first-order Markov dynamics encoded in $p(x_t \mid x_{t-1}, \theta)$
- Observations are generated from hidden states according to $p(y_t \mid x_t, \theta)$



State-Space Models (SSM)

Example for T=1**:**
Likelihood of **complete data:**

$$p(x_1, y_1 \mid \theta) = p(x_1 \mid x_0, \theta)\, p(y_1 \mid x_1, \theta)$$

Likelihood of **incomplete** data:

$$p(y_1 \mid \theta) = \int p(x_1 \mid x_0, \theta)\, p(y_1 \mid x_1, \theta)\, dx_1$$

This is a continuous mixture model:

$$p(y_1 \mid \theta) = \int f(\omega)\, p(y_1 \mid \omega, \theta)\, d\omega$$

For any **fixed** time interval t, y_t has a continuous mixture distribution, but a SSM links these distributions for different values of t into a single **process** for $\left(x_t, y_t\right)$

State-Space Models (SSM)

Observable N-dimensional data: $y_{1:T} = y^{(1)}, y^{(2)}, \ldots, y^{(T)}$
Hidden state sequence: $x_{1:T} = x^{(1)}, x^{(2)}, \ldots, x^{(T)}$

The dynamics is first-order Markov in the **continuous** hidden state:

$$p(x_{1:T}, y_{1:T} \mid \theta) = \prod_{t=1}^{T} p(x_t \mid x_{t-1}, \theta)\, p(y_t \mid x_t, \theta)$$

- Hidden states x have first-order Markov dynamics encoded in $p(x_t \mid x_{t-1}, \theta)$
- Observations are generated from hidden states according to $p(y_t \mid x_t, \theta)$

Two computational problems:

Forecasting future value $p(y_{t+1} \mid x_{1:t}, y_t, \theta) = Inference + Learning$

1. **Inference** of the hidden state $p(x_{1:t} \mid y_{1:t}, \theta)$
2. **Learning** of model parameters θ

Linear Gaussian State-Space Model

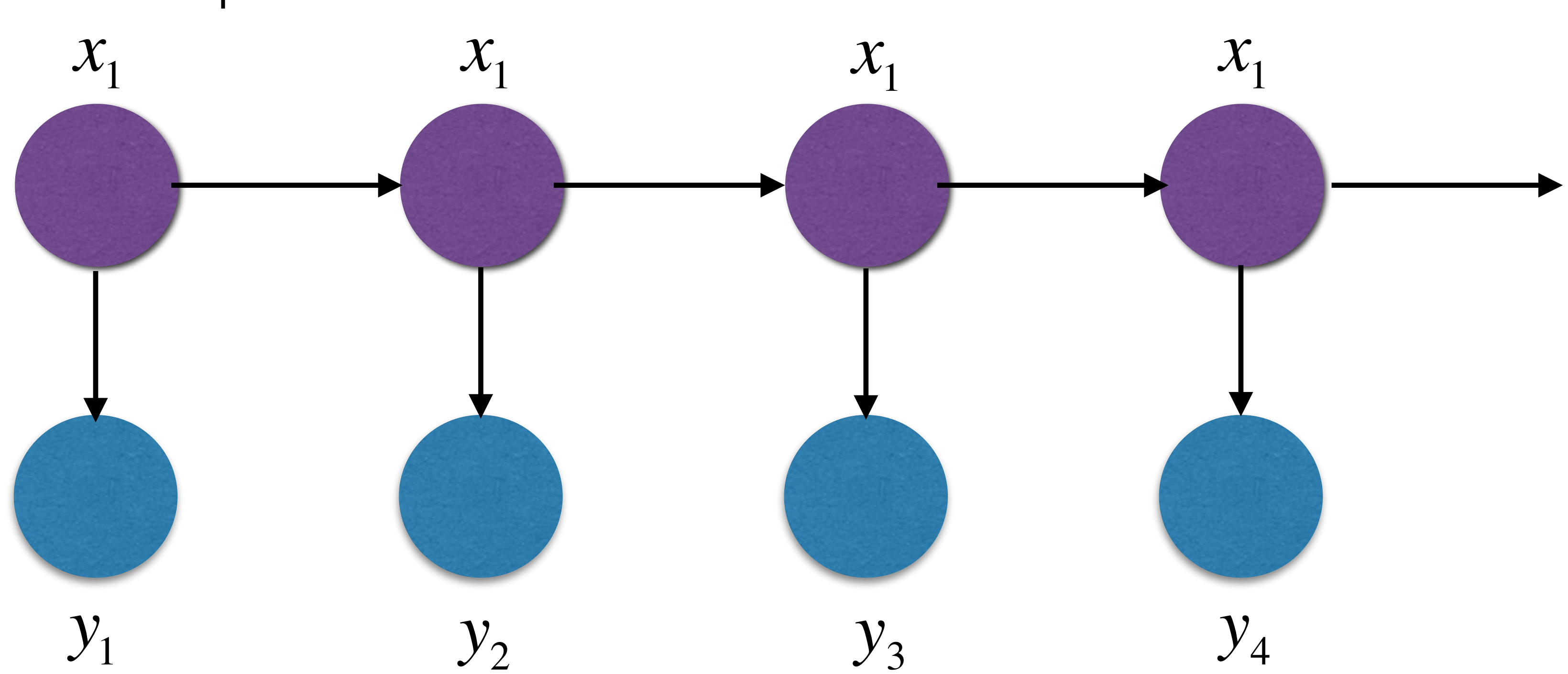
The linear-Gaussian realization of dynamics is first-order Markov in the hidden state $x^{(t)}$:

$$p(x_{1:T}, y_{1:T} \mid \theta) = \prod_{t=1}^{T} p(x_t \mid x_{t-1}, \theta)\, p(y_t \mid x_t, \theta)$$

$$x_t = Ax_{t-1} + w_t \;\Leftrightarrow\; p(x_t \mid x_{t-1}, \theta) = N(Ax_{t-1}, \sigma_w^2)$$

$$y_t = Cx_t + v_t \;\Leftrightarrow\; p(y_t \mid x_t, \theta) = N(Cx_{t-1}, \sigma_v^2)$$

- The linear-Gaussian SSM is a time-series generalization of Factor Analysis
- Inference and learning can be done using the EM algorithm: E-step: learn the posterior over hidden variables $p(x_{1:T} \mid y_{1:T}, \theta)$ (*Kalman smoothing* algorithm), M-step: estimate all parameters for a fixed distribution of hidden variables.



State-space models in Finance

- The latent variable approach to partially observable firm capital structure (M. Roberts, “The Dynamics of Capital Structure: An Empirical Analysis of a Partially Observed System”,2002)
- Problem studied: adjustment of firm leverage=debt/(debt+equity) to target values
- Hidden factors: Gaussian hidden factors describing the unobserved “true” values of marginal tax rate, probability of bankruptcy, firm size, investment opportunities and average industry leverage
- All hidden factors are modeling as auto-regression AR(1) processes
- Observed factors and leverage are obtained as hidden variables plus Gaussian noise
- Formulated and estimated as a state-space model
- A state-space model approach allows one to handle missing data
- Results may be used for other purposes, for example to improve default predictions (G. Loeffler and A. Mauler, “Incorporating the dynamics of leverage into default prediction”, 2009).

Control question

Select all correct answers

1. State-Space models specify a first-order Markov dynamics for the observed signal $y^{(t)}$.
2. State-Space have two sets of unknowns - a hidden state and model parameters.
3. State-Space have two sets of unknowns: variables describing the state and the space location of all data points, respectively.
4. Linear-Gaussian State-Space Models are obtained if the state equation is linear, and both the state and observational noise are Gaussian
5. Inference and Learning in State-Space models can be done using the EM algorithm.
6. When variances of both the state noise and observational noise are set to zero, State-Space models become Hidden Markov Models (HMM)

Correct answers: 2, 4, 5