Algorithms for context prediction in Ubiquitous Systems

Prediction by stochastic approaches

Stephan Sigg

Institute of Distributed and Ubiquitous Systems
Technische Universität Braunschweig

February 3, 2009

Overview and Structure

- Introduction to context aware computing
- Basics of probability theory
- Algorithms
    - Simple prediction approaches
    - Markov prediction approaches
    - The State predictor
    - Alignment prediction
    - Prediction with self organising maps
    - Stochastic prediction approaches
    - Prediction by random search heuristics

Overview and Structure

- Introduction to context aware computing
- Basics of probability theory
- **Algorithms**
  - Simple prediction approaches
  - Markov prediction approaches
  - The State predictor
  - Alignment prediction
  - Prediction with self organising maps
  - **Stochastic prediction approaches**
  - Prediction by random search heuristics

Outline

Stochastic prediction approaches

1. Introduction
2. Time series prediction models
   - Moving average models
   - Autoregressive models
   - ARMA models
3. Properties of stochastic time series prediction approaches

Introduction
Historical remarks

- Despite recent developments with nonlinear models, some of the most common stochastic models in time series prediction are parametric linear models as autoregressive (AR), moving average (MA) or autoregressive moving average (ARMA) processes[1]
- Examples for application scenarios:
  - Financial time series prediction
  - Wind power prediction.

[1] Jens-Peter KreiÃŸ and Georg Neuhaus, *EinfÃ¼hrung in die Zeitreihenanalyse, Springer, 2006.*

Introduction

General model

- Assume a stochastic process $\pi(t)$ that generates outputs $\chi(t)$ at each point t in time
  - Random values $\chi(t)$ can be univariate or multivariate and can take discrete or continuous values
  - Time can also be either discrete or continuous.
- Task: Find parameters $\Theta = \{\theta_1, \ldots, \theta_n\}$ that describe the stochastic mechanism[2]
- Prediction accomplished by calculating conditional probability density $P(\chi(t)|\{\Theta, \{\chi(t-1), \ldots, \chi(t-m)\}\})$.

[2] R.O. Duda, P.E. Hart and D.G. Stork, *Pattern Classification*, Wiley Interscience, 2001.

Outline

Stochastic prediction approaches

1. Introduction

2. Time series prediction models
   - Moving average models
   - Autoregressive models
   - ARMA models

3. Properties of stochastic time series prediction approaches

Time series prediction models

Moving average models

Moving average (MA) models

- Let $Z(t)$ be some fixed zero-mean, unit-variance random process.
- $\chi(t)$ is a $MA(k)$ process (MA-process of order k), if

$$\chi(t) = \sum_{\tau=0}^{k} \beta_\tau Z(t-\tau) \quad (1)$$

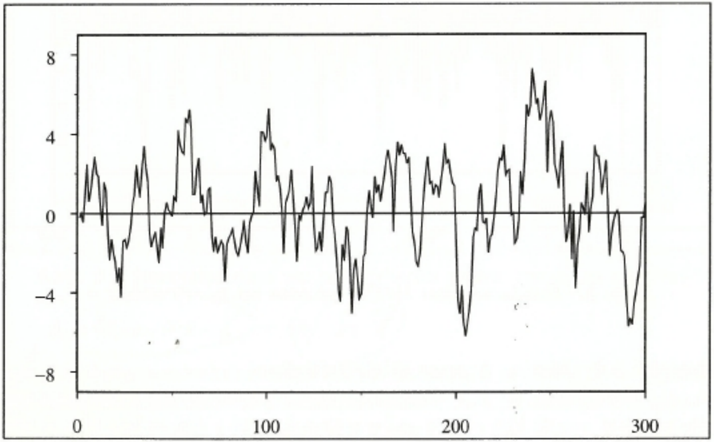
where the β_τ are constants.

- Moving average processes are utilised to describe stochastic processes with finite, short-term linear memory [3]

[3] C. Chatfild, *The Analysis of Time Series: An Introduction*, Chapman and Hall, 1996.

Time series prediction models

Moving average models



Time series prediction models

Autoregressive models

Autoregressive (AR) models

- AR processes, the values at time t depend linearly on previous values
- $\chi(t)$ is an $AR(k)$ process of order k, if

$$\sum_{\nu=0}^{k} \alpha_\nu \chi(t-\nu) = Z(t) \tag{2}$$

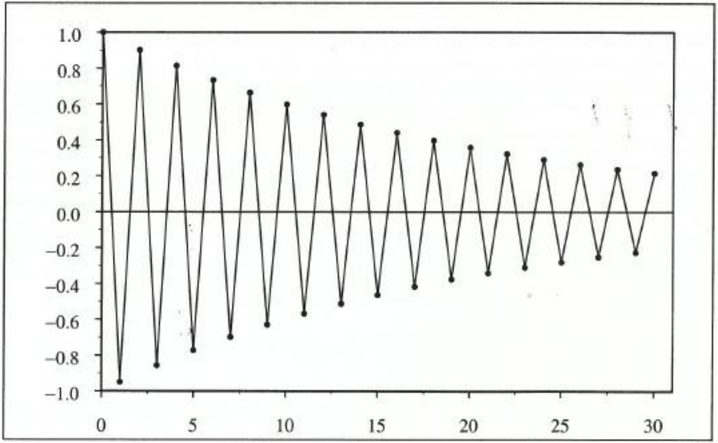
where α_ν are constants.

- Autoregressive processes are used to capture exponential traces [4]

[4] C. Chatfild, *The Analysis of Time Series: An Introduction*, Chapman and Hall, 1996.

Time series prediction models

Moving average models



Time series prediction models
ARMA models

ARMA models

- ARMA processes are a combination of AR and MA processes.
- An $ARMA(p,q)$ process is a stochastic process $\chi(t)$ in which

$$\sum_{\nu=0}^{p} \alpha_\nu \chi(t-\nu) = \sum_{\tau=0}^{q} \beta_\tau Z(t-\tau) \qquad (3)$$

, where $\{\alpha_\nu, \beta_\tau\}$ are constants [a]

[a] C. Chatfild, *The Analysis of Time Series: An Introduction*, Chapman and Hall, 1996.

Time series prediction models

Context prediction with ARMA approaches

- ARMA process is already designed to approximate a numeric time series in time
- Only requirement for context prediction: Numerical context types

Time series prediction models

ARMA models – Discussion

- ARMA methods provide a powerful tool to approximate stochastic processes
- ARMA processes are able to achieve excellent results in context prediction scenarios [5]
- Applicable to one-dimensional, as well as multi-dimensional, data sets
- Computational complexity can be estimated as $O(k \log(k))$ [6]
- No prior pre-processing or separate learning tool required.
- Only applicable to contexts of numeric context data type.

[5] Rene Michael Mayrhofer, *An Architecture for Context Prediction*, Johannes Kepeler University of Linz, 2004.

[6] J. Cadzow and K. Ogino, *Adaptive ARMA spectral estimation* Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 1981.

Outline

Stochastic prediction approaches

1. Introduction
2. Time series prediction models
   - Moving average models
   - Autoregressive models
   - ARMA models
3. Properties of stochastic time series prediction approaches

Aspects of prediction algorithms

Summary

	IPAM	ONISI	Markov	CRF	ARMA
Numeric Contexts	yes	no	no	no	yes
Non-numeric Contexts	yes	yes	yes	yes	no
Complexity	$O(k)$	$O(k^2)$	$O(C^2)$	$O(C^2)$	$O(k \log(k))$
Learning ability	(no)	yes	yes	yes	(no)
Approximate matching	no	no	no	no	(no)
Multi-dim. TS	(no)	(no)	(no)	(no)	yes
Discrete data	yes	yes	yes	yes	yes
Variable length patterns	no	yes	no	(yes)	yes
Multi-type TS	yes	no	(no)	(no)	no
Continuous data	no	no	no	no	yes
Pre-processing	$O(k)$	–	$O(k)$	$O(k)$	–
Context durations	no	no	no	no	yes
Continuous time	no	no	yes	yes	yes

Aspects of prediction algorithms

Summary

	SPM	Align	SOM	EA
Numeric Contexts	yes	yes	yes	yes
Non-numeric Contexts	yes	yes	yes	yes
Complexity	$O(1)$	$O(l \cdot k^2)$	$O(\log(\|k\|))$	$O(\log(\|k\|))$
Learning ability	(yes)	yes	yes	yes
Approximate matching	no	yes	yes	yes
Multi-dim. TS	(no)	yes	yes	yes
Discrete data	yes	yes	yes	yes
Variable length patterns	yes	yes	yes	yes
Multi-type TS	no	yes	yes	yes
Continuous data	no	no	no	no
Pre-processing	$O(k)$	$O(k^2)$	$O(k^2)$	–[7]
Context durations	no	yes	yes	yes
Continuous time	no	no	no	no

[7] Preprocessing complexity dependent on problem definition and algorithmic modelling. Presumably high preprocessing required for random search approach