Using Linguistic Information in Density Estimation

U. Kaymak
Erasmus University Rotterdam
Erasmus School of Economics, Econometric Institute
u.kaymak@ieee.org

LFA 2008, Lens, France
16-17 October 2008

Erasmus University Rotterdam

Summary

Goal: discuss probabilistic fuzzy systems to estimate density functions constrained by linguistic information

Outline:

- Introduction
- Fuzzy models and interpretability
- Probabilistic Fuzzy Systems (PFS)
  - Mamdani PFS
  - Takagi-Sugeno PFS
- Parameter estimation in PFS
- Application examples
  - VaR estimation
  - Modelling of conditional returns in financial markets
- Conclusions

Contributors

- Jan van den Berg
- Willem-Max van den Bergh
- Ludo Waltman
- Rui J. Almeida
- Du Xu

Fuzzy systems

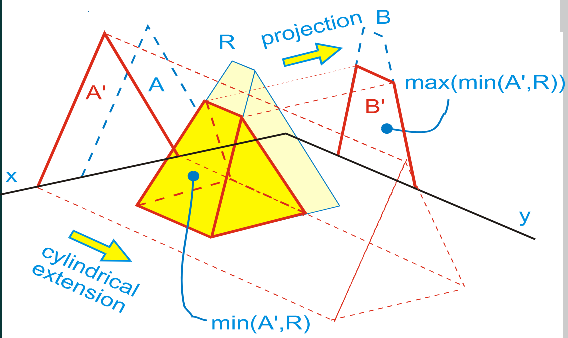
Many successful applications and solutions

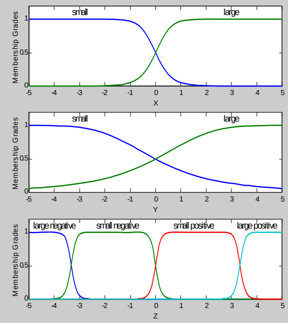
- Data analysis and modelling
- Image processing
- Control systems
- Household appliances
- Consumer electronics
- Multi-agent system design
- Decision making under uncertainty
- Information retrieval

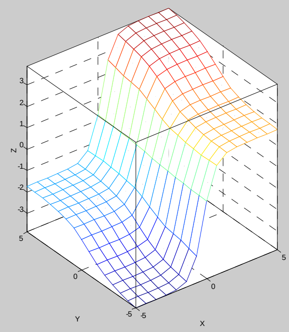
Erasmus University Rotterdam

Mamdani fuzzy models

- Fuzzy antecedents,
  fuzzy consequents
  if x is A and y is B then z is C
- Compositional rule of inference







Erasmus University Rotterdam

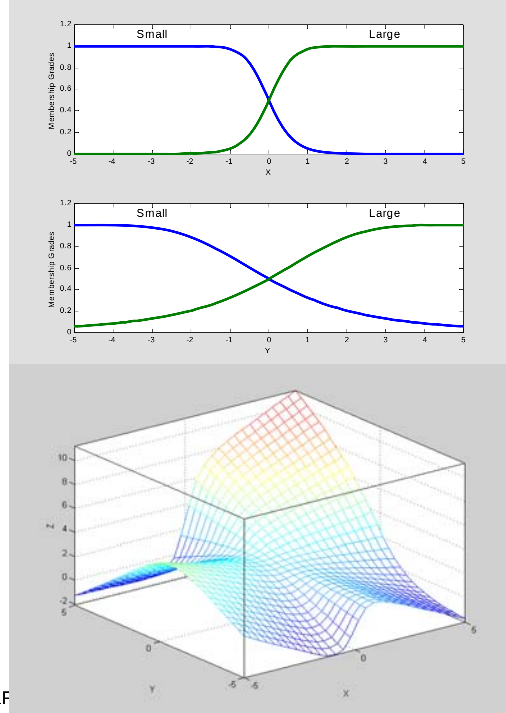
Takagi-Sugeno fuzzy models

- Fuzzy antecedents, crisp consequents
- Consequent is a crisp function of inputs
  if x is A and y is B then z = f(x,y)
- Zero-order Sugeno: constant consequent
  if x is A and y is B then z = c
- First-order Sugeno: linear consequent
  if x is A and y is B then z = ax+by+c
- Overall output is a weighted average of individual rule outputs

$$z^* = \frac{\sum_{q=1}^{Q} w_q z_q}{\sum_{q=1}^{Q} w_q}$$

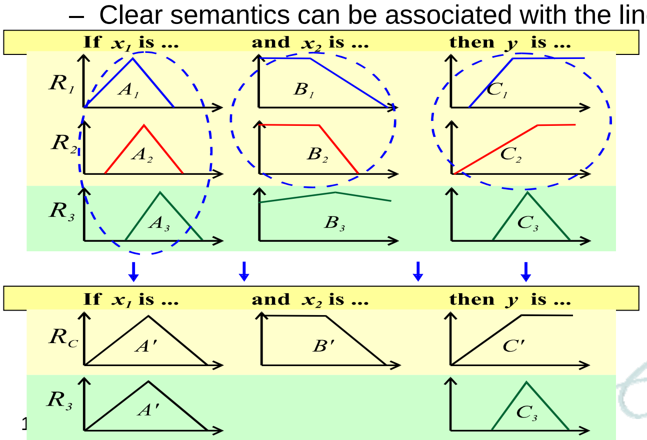
1st order Sugeno double input

- If X is Small and Y is Small then z = -x+y+1
- If X is Small and Y is Large then z = -y+3
- If X is Large and Y is Small then z = -x+3
- If X is Large and Y is Large then z = x+y+2



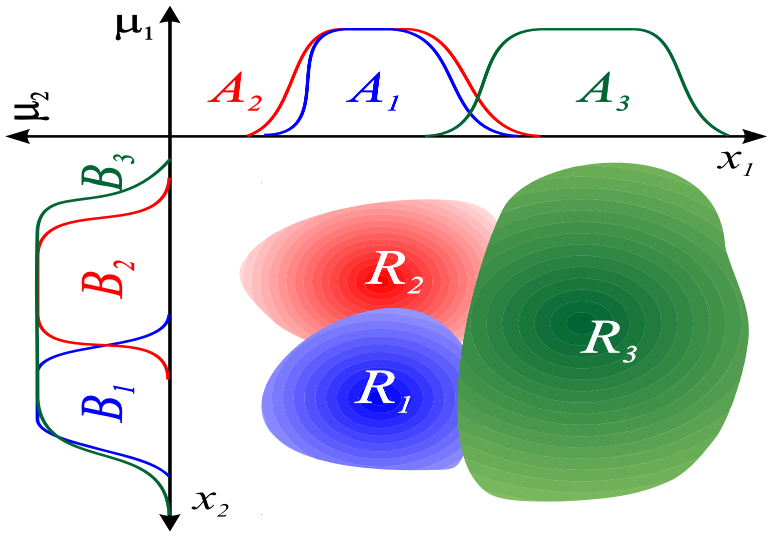
Interpretability of fuzzy models

- Design a model s.t.
  - Model complexity is small (nr. of rules and linguistic terms)
  - Linguistic terms and rules are sufficiently distinct
  - Clear semantics can be associated with the linguistic terms



Erasmus University Rotterdam

Redundancy after clustering



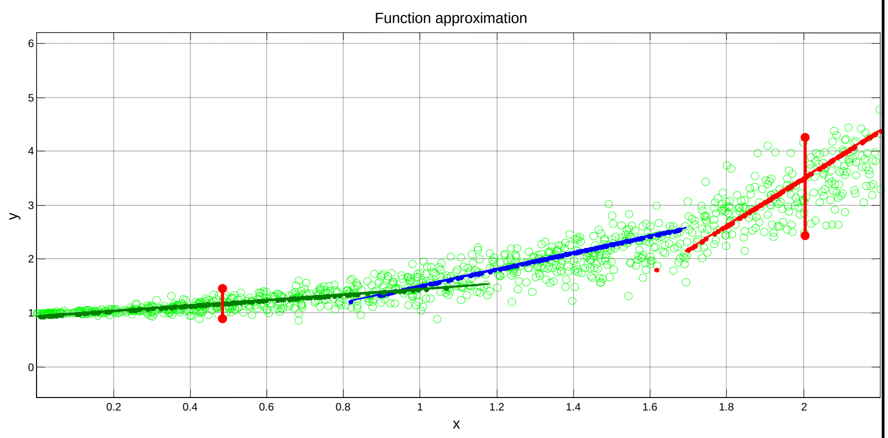
Erasmus University Rotterdam

Similarity based simplification

Reduce redundancy in rule and term set to improve transparency

- Merge similar antecedent fuzzy sets
  - create generalized concepts
  - reduce the number of terms
- Remove sets similar to universal set (always fires)
  - reduce number of terms
- Combine / merge similar consequents
  - reduce the number of consequent values
- Combine rules with equal antecedents
  - reduce number of rules

Deterministic output, probabilistic uncertainty



Density estimation and fuzzy systems

- After defuzzification, fuzzy systems implement a nonlinear mapping from inputs to outputs
- Output is assumed to be deterministic: regular fuzzy systems cannot estimate densities
- Many problems require estimation of densities: extend fuzzy systems to estimate density
- In doing so, linguistically interpretable models can also be developed for probability density estimation
  → probabilistic fuzzy systems

Applications of density estimation

- Computing error bounds on models
- Value-at-risk estimation for financial risks
- Linguistic descriptions of probability distributions
- Ore-grade estimations
- Linguistic descriptions for stochastic time series

Erasmus University Rotterdam

16-Oct-2008 LFA 2008

Probabilistic vs. fuzzy systems

Probabilistic systems

- Consider uncertainty as randomness
- Emphasis on statistical properties of data
- Axiomatic grounding
- Assumptions often taken as a priori

Fuzzy systems

- Emphasis on linguistic uncertainty
- Statistical properties of data often ignored
- Function approximation properties (deterministic)
- Focus on linguistic grounding

Erasmus University Rotterdam

16-Oct-2008 LFA 2008

Hammer principle

If your only tool is a hammer, you start seeing the whole world as a nail
(H.-J. Zimmermann)

Preliminaries

Probability of a fuzzy event:

- Crisp

$$\Pr(A) = \int_{x \in A} p(x)\,dx = \int_X \chi_A(x) p(x)\,dx$$

- Fuzzy (Zadeh, 1968)

$$\Pr(A) = \int_X \mu_A(x) p(x)\,dx$$

Probability of a fuzzy event is the expectation of its membership function

Conditional probability:

$$\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\int_X \mu_A(x)\mu_B(x)p(x)\,dx}{\int_X \mu_B(x)p(x)\,dx}$$

Simple estimation

- Let $x_1, \ldots, x_n$ be a random sample on a domain X
- The probability of a crisp event A can be estimated by

$$\frac{1}{n}\sum_k \chi_A(\mathbf{x}_k)$$

- The probability of a fuzzy event A_i can be estimated by

$$\frac{1}{n}\sum_k \mu_{A_i}(\mathbf{x}_k)$$

assuming that X is well-formed, i.e. $\forall \mathbf{x}_k \sum_i \mu_{A_i}(\mathbf{x}_k) = 1$

Deterministic and probabilistic rules

Rule R_q : If $\mathbf{x}$ is A_q then $\underline{y}$ is C_q

Model parameters: A_q, C_q

If current returns are large,
then future returns will be large

If current returns are large,
then future returns will be large with probability ρ_1
future returns will be small with probability $1-\rho_1$

Probabilistic fuzzy rules

Rule R_q : If $\mathbf{x}$ is A_q then

$$\underline{y} \text{ is } C_{1q} \text{ with } \Pr(C_{1q}|A_q) \text{ and ... and}$$
$$\underline{y} \text{ is } C_{jq} \text{ with } \Pr(C_{jq}|A_q) \text{ and ... and}$$
$$\underline{y} \text{ is } C_{Nq} \text{ with } \Pr(C_{Nq}|A_q)$$

Model parameters: $A_q, C_{jq}, \Pr(C_{jq}|A_q)$

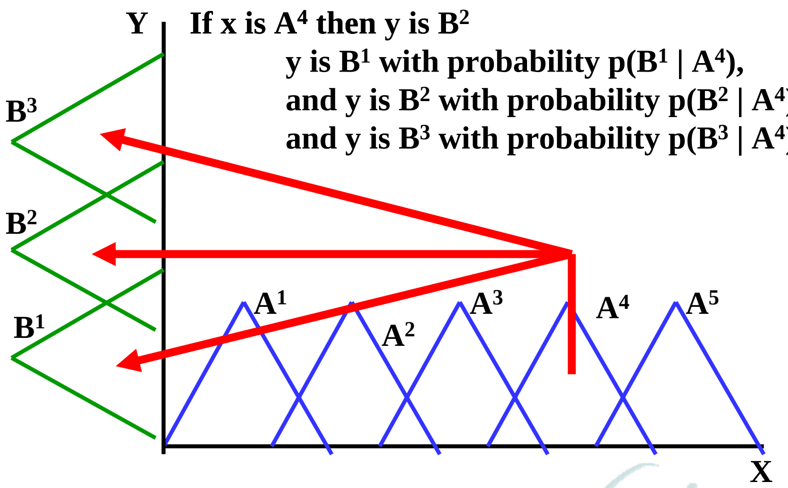
For practical purposes, one can assume that

$$\forall j,q,q' \quad C_{jq} = C_{jq'}$$

Deterministic vs. probabilistic FS

If x is A^4 then y is B^2

y is B^1 with probability $p(B^1 | A^4)$,
and y is B^2 with probability $p(B^2 | A^4)$,
and y is B^3 with probability $p(B^3 | A^4)$.



Probability of a fuzzy statement

What is the probability that a randomly selected French person is very tall?

- Probability is unlikely (low, small, etc.)
- Probability is about 0.4
- Probability is 0.422 ⟵ **PFS approach**

Probabilistic fuzzy system

$$\text{Rule } R_q : \text{If } \mathbf{x} \text{ is } A_q \text{ then } p(y) = p(y \mid A_q)$$

Additive reasoning:

$$p(y \mid \mathbf{x}) = \frac{\sum_{q=1}^{Q} \mu_{A_q}(\mathbf{x})\, p(y \mid A_q)}{\sum_{q=1}^{Q} \mu_{A_q}(\mathbf{x})} = \sum_{q=1}^{Q} \beta_q(\mathbf{x})\, p(y \mid A_q)$$

$$y = \mathrm{E}(y \mid \mathbf{x}) = \sum_{q=1}^{Q} \beta_q(\mathbf{x})\, \mathrm{E}(y \mid A_q)$$

Probabilistic Mamdani systems

Rule R_q : If $\mathbf{x}$ is A_q then

$$\underline{y} = C_1 \text{ with } \Pr(C_1|A_q) \text{ and ... and}$$
$$\underline{y} = C_j \text{ with } \Pr(C_j|A_q) \text{ and ... and}$$
$$\underline{y} = C_N \text{ with } \Pr(C_N|A_q)$$

Reasoning: $y = \mathrm{E}(y\,|\,\mathbf{x}) = \sum_{q=1}^{Q}\sum_{j=1}^{N} \beta_q(\mathbf{x}) \Pr(C_j|A_q)\, z_j$

Centroid of fuzzy consequent set $z_j = \dfrac{\int_Y y\mu_{C_j}(y)dy}{\int_Y \mu_{C_j}(y)dy}$

Probabilistic TS systems

Zero-order probabilistic Takagi-Sugeno

Rule R_q : If $\mathbf{x}$ is A_q then

$$y = y_1 \text{ with } \Pr(y_1|A_q) \text{ and}$$
$$y = y_2 \text{ with } \Pr(y_2|A_q) \text{ and ... and}$$
$$y = y_N \text{ with } \Pr(y_N|A_q)$$

In essence, consequents are crisp sets centered around y_j

Reasoning: $y = \mathrm{E}(y\,|\,\mathbf{x}) = \sum_{q=1}^{Q}\sum_{j=1}^{N} \beta_q(\mathbf{x}) y_j \Pr(y_j|A_q)$

Constrained consequent partition

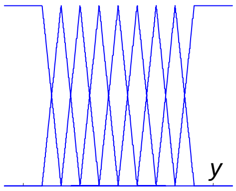
If $\mathbf{x}$ is A_q then

y is C_{q1} with $\Pr(C_{q1}|A_q)$ and

y is C_{q2} with $\Pr(C_{q2}|A_q)$ and ...

y is C_{qN} with $\Pr(C_{qN}|A_q)$.

$$\sum_{j=1}^{N} \mu_{C_j}(y) = 1$$



$$f(y|\mathbf{x}) = \sum_{j=1}^{N} \frac{\sum_{q=1}^{Q} \beta_q(\mathbf{x}) \Pr(C_j|A_q) \mu_{C_j}(y)}{\int_{-\infty}^{\infty} \mu_{C_j}(y) dy}$$

Histograms

- Let $x_1, \ldots, x_n$ be a random sample from a univariate distribution with pdf $f(x)$
- Let the characteristic functions $\chi_i(x)$ (defining crisp bins/intervals A_i) constitute a *crisp partitioning:*

$$\forall x \in \Re : \sum_i \chi_i(x) = 1 \;\wedge\; \forall i \neq j : \chi_i(x).\chi_j(x) = 0$$

- A histogram estimates $f(x)$ as follows:

$$\hat{f}(x) = \sum_i \frac{\chi_i(x) p_i}{|A_i|} \approx \sum_i \frac{\chi_i(x)\left(\frac{1}{n}\sum_k \chi_i(x_k)\right)}{\int \chi_i(x) dx}$$

Fuzzy histograms

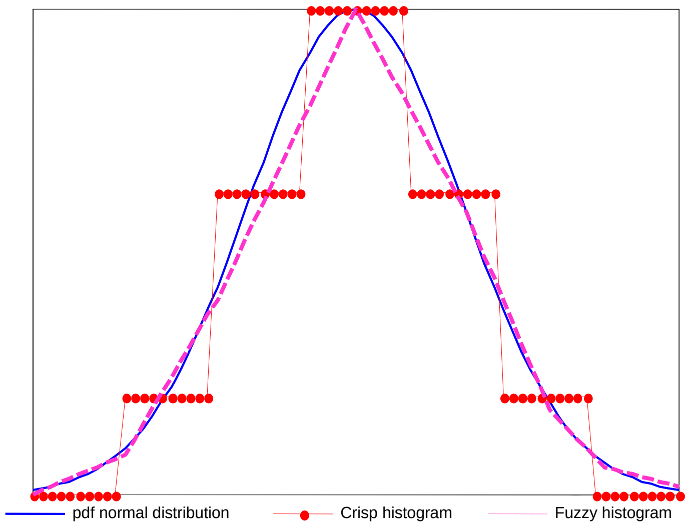
- Let $x_1, \ldots, x_n$ be a *random sample* of *size n* from a univariate distribution with pdf $f(x)$
- Let the membership functions functions $\mu_i(x)$ (defining fuzzy bins A_i) constitute a *fuzzy partitioning:*

$$\forall x \in \Re : \sum_i \mu_i(x) = 1$$

- A fuzzy histogram estimates $f(x)$ as follows:

$$\hat{f}(x) = \sum_i \frac{\mu_i(x) p_i}{|A_i|} \approx \sum_i \frac{\mu_i(x)\left(\frac{1}{n}\sum_k \mu_i(x_k)\right)}{\int \mu_i(x)dx}$$

Crisp vs. fuzzy histogram



Two interpretations

- Original interpretation:

$$\hat{f}(x) = \sum_i \frac{\mu_i(x)\, p_i}{\int \mu_i(x)dx} = \sum_i \mu_i(x) \frac{p_i}{\int \mu_i(x)dx}$$

concerns an interpolation between local densities

- Another interpretation:

$$\hat{f}(x) = \sum_i \frac{\mu_i(x)\, p_i}{\int \mu_i(x)dx} = \sum_i p_i \frac{\mu_i(x)}{\int \mu_i(x)dx}$$

concerns a sum of pdfs, weighted with p_i

Erasmus University Rotterdam

Relation to deterministic FSs

- Zero-order Takagi-Sugeno system

Takagi-Sugeno reasoning $y^* = \sum_{q=1}^{Q} \beta_q(\mathbf{x}) c_q$

c.f. $y = \mathrm{E}(y \mid \mathbf{x}) = \sum_{q=1}^{Q} \sum_{j=1}^{N} \beta_q(\mathbf{x}) \Pr(y_j | A_q)\, z_j$

Select $c_q = \sum_{j=1}^{N} \Pr(y_j | A_q)\, z_j$

Erasmus University Rotterdam

summary

Probabilistic fuzzy systems

- Essentially a fuzzy system that estimates a probability density function, i.e. the fuzzy system approximates a p.d.f.
- Usually p.d.f. is conditional on the input
- Linguistic information is coded in fuzzy rules
- Related to density estimation techniques such as Parzen windows and kernel based density estimation
- Combine linguistic uncertainty with probabilistic uncertainty
- Different types of fuzzy systems can be extended to the PFS equivalent (e.g. Mamdani fuzzy systems, Takagi-Sugeno fuzzy systems)

PFS design

- Identifying mental world vs. observed world (van den Eijkel 1999)
- Mental world: linguistic descriptions, fuzzy conceptualization, experts' knowledge
- Observed world: data measurements, probability density functions, optimal consequent parameters
- Optimal design given a mental world: application of conditional probability measures for fuzzy events
- Optimal design given an observed world: nonlinear optimization techniques

Implications for adaptation/learning

- Adaptation/learning required for two fundamentally different quantities
- Tracking of changes in mental models as well as observed data is required
- In general, more flexibility through more parameters
- Suited for different types of adaptation/learning

Parameter estimation

- Parameters
  - Antecedent membership functions
  - Consequent membership functions
  - Probability parameters of rules
- Estimation methods
  - Joint estimation: estimate all parameters simultaneously (complex optimisation, local optimum)
  - Sequential estimation: estimate membership functions first (expert driven or data driven) and then estimate probability parameters (membership weighted counting or maximum likelihood)

Parameter estimation using a sequential method

- First the antecedent membership functions $\mu_{A_q}(\mathbf{x})$ are estimated. This can be done using, for example, fuzzy c-means clustering.
- Then the probability parameters $\Pr(C_j|A_q)$ are estimated by setting the probability parameters equal to estimates of conditional probabilities.

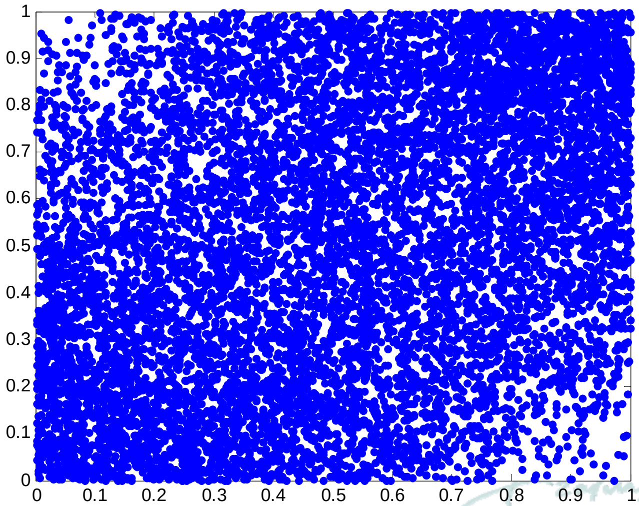
Estimation of probability parameters

- Conditional probabilities $\Pr(C_j \mid A_q)$ can be assessed directly by using the definition of the probability of joint events:

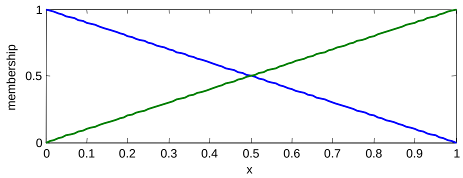
$$\Pr(C_j \mid A_q) = \frac{\Pr(A_q \cap C_j)}{\Pr(A_q)} \approx \frac{\sum_{(\mathbf{x}_k, y_k)} \mu_{A_q}(\mathbf{x}_k)\mu_{C_j}(y_k)}{\sum_{\mathbf{x}_k} \mu_{A_q}(\mathbf{x}_k)}$$

- This method does not provide maximum likelihood estimates of the probability parameters.

Bias in parameter estimation

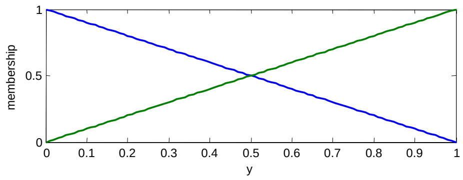


Underlying data generation process



$$\mu_{A_1}(x) = 1 - x$$

$$\mu_{A_2}(x) = x$$



$$\mu_{c_1}(y) = 1 - y$$

$$\mu_{C_2}(y) = y$$

$$\Pr(y \mid x) = 4xy - 2x - 2y + 2$$

Optimal parameters: $\Pr(C_1 \mid A_1) = 1$ $\Pr(C_1 \mid A_2) = 0$

$\Pr(C_2 \mid A_1) = 0$ $\Pr(C_2 \mid A_2) = 1$

Parameter estimation using maximum likelihood (1)

- Both the antecedent membership functions $\mu_{A_q}(\mathbf{x})$ and the probability parameters p_{jq} are estimated by maximizing the likelihood of the training examples, which is given by

$$L = \prod_{k=1}^{n} \hat{p}(y_k \mid \mathbf{x}_k).$$

- This is equivalent to minimization of the negative log-likelihood

$$-\sum_{k=1}^{n} \ln \hat{p}(y_k \mid \mathbf{x}_k).$$

Parameter estimation using maximum likelihood (2)

- The optimization problem is constrained by

$$\forall j, q : p_{jq} \geq 0$$

$$\forall q : \sum_{j=1}^{N} p_{jq} = 1.$$

- An unconstrained optimization problem can be obtained by using auxiliary variables u_{jq} that are related to the probability parameters p_{jq} according to the softmax function

$$p_{jq} = \frac{e^{u_{jq}}}{\sum_{j'=1}^{N} e^{u_{j'q}}}$$

Experimental comparison (1)

- Use Gaussian membership functions

$$\mu_{A_q}(\mathbf{x}) = \exp\left(-\sum_{l=1}^{d} \frac{(x_l - c_{ql})^2}{\sigma_{ql}^2} \right)$$

- The centers c_{ql} are determined using fuzzy c-means clustering

- The widths σ_{ql} are set equal to $\sigma_{ql} = \min_{j' \neq j} ||\mathbf{c}_q - \mathbf{c}_{q'}||$

Experimental comparison (2)

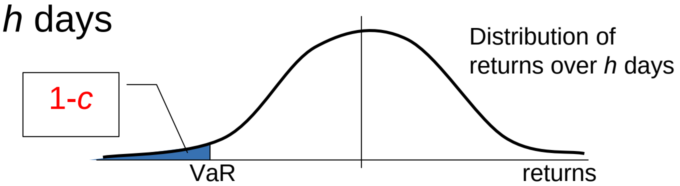
- Misclassification rates

	Wisconsin breast cancer	Wine
Sequential method	0.261 (0.036)	0.034 (0.048)
Maximum likelihood	0.029 (0.021)	0.023 (0.041)

- Calculated using ten-fold cross-validation
- Standard deviations reported within parentheses

Value-at-risk

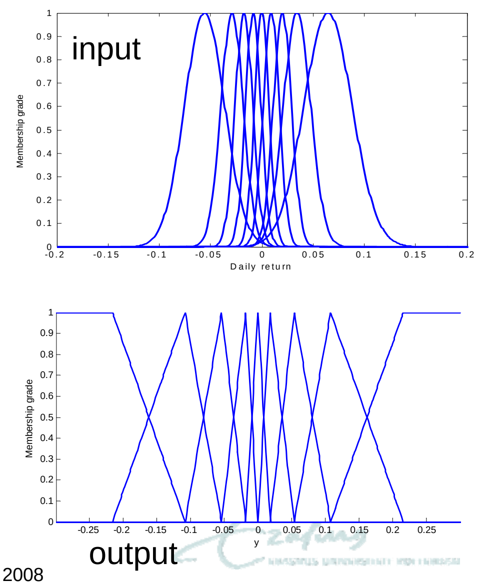
- Quantifies downside risk (risk of making losses in financial markets)
- Single number for senior management
- Indicates maximum loss with certainty *c* that a portfolio of assets might suffer over a horizon of *h* days



- Most VaR models reduce to estimating the volatility of the returns of a portfolio

Probabilistic fuzzy VaR model

- One day estimations for returns (h = 1)
  - Input: yesterday's returns, output: today's returns
- Consider return distributions of the total portfolio
- Output membership functions defined by the modeller
- Input membership functions determined by FCM clustering and replacing the clusters by Gaussian membership functions
- Probability parameters determined by maximum likelihood estimation
- Output membership functions are scaled by a data-dependent value *z*



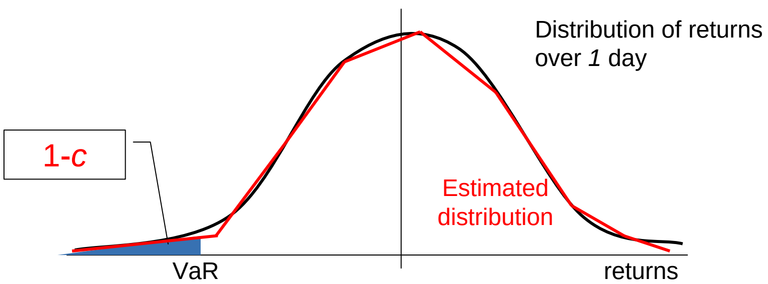
Erasmus University Rotterdam

Modelling algorithm

1. Collect price series; compute one-period returns; create training and test data
2. Determine antecedent membership functions: apply fuzzy c-means clustering to compute the locations of the membership functions and use cluster covariance to obtain the spreads
3. Select the number of consequent membership functions and form an output partition; determine the scaling factor *z*.
4. Determine optimal probability parameters by maximum likelihood
5. Compute the estimated conditional probability distribution function for the one-period returns for each observation in the test set
6. Given the conditional probability distribution functions, compute the VaR
7. Validate the model

Erasmus University Rotterdam

Density estimation and approximation



Experimental study

- Daily returns for six large companies from Dow Jones and Shanghai Stock Exchange
- Data collected for 1000 days
- Training data for first 500 days
- Optimal value of parameter *z* determined by simple search for each data set
- Comparison with a GARCH(1,1) model

STOCKS USED IN THE EMPIRICAL STUDY.

	Data Range
KPN	06/01/1999 – 27/12/2002
ABN	05/01/2000 – 29/12/2003
JiaLing	23/04/2002 – 17/07/2006
BaoShan	04/01/2001 – 04/04/2005
COSCO	26/09/2002 – 16/07/2007
Merchant Bank	10/05/2002 - 28/08/2006

Erasmus University Rotterdam

16-Oct-2008

Probability parameters

initial

	Consequent								
Rule	1	2	3	4	5	6	7	8	9
1	0.1003	0.1333	0.2066	0.0567	0.0441	0.1215	0.1012	0.1218	0.1143
2	0.0565	0.1351	0.1659	0.0792	0.0744	0.0972	0.1617	0.1489	0.0811
3	0.0459	0.1679	0.1495	0.1300	0.1024	0.0773	0.1463	0.1359	0.0448
4	0.0544	0.1683	0.1700	0.1105	0.0647	0.0802	0.1454	0.1579	0.0468
5	0.0516	0.1578	0.1800	0.1119	0.0692	0.0770	0.1547	0.1472	0.0506
6	0.0563	0.1648	0.1760	0.0956	0.0655	0.1206	0.1340	0.1227	0.0646
7	0.0700	0.1877	0.1659	0.0748	0.0405	0.0901	0.1147	0.2002	0.0562
8	0.0529	0.1625	0.1626	0.1122	0.0999	0.0930	0.1313	0.1286	0.0570
9	0.0539	0.1729	0.1624	0.1132	0.0746	0.0772	0.1476	0.1505	0.0476

optimised

	Consequent								
Rule	1	2	3	4	5	6	7	8	9
1	0.1401	0.0850	0.3776	0	0	0.1586	0	0.0940	0.1446
2	0.0247	0.0762	0	0	0.1366	0.1239	0.3359	0.2380	0.0646
3	0	0.0308	0	0.3430	0.2831	0	0.3432	0	0
4	0.0017	0.1197	0.0895	0.1036	0.0348	0.1577	0.1609	0.3321	0
5	0	0.1893	0.4046	0.2158	0.0725	0	0.1178	0	0
6	0	0	0.2309	0.1297	0	0.3170	0.2131	0	0.1093
7	0	0.4398	0.1626	0.0046	0	0	0	0.3931	0
8	0.0455	0.3267	0.2359	0	0.0755	0.0511	0.0120	0.2105	0.0429
9	0.0595	0.3158	0.2772	0	0	0.0478	0.0980	0.1970	0.0047

Erasmus University Rotterdam

Back testing results

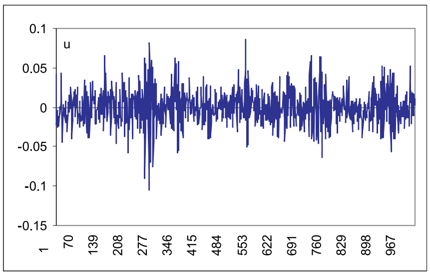
Erasmus University Rotterdam

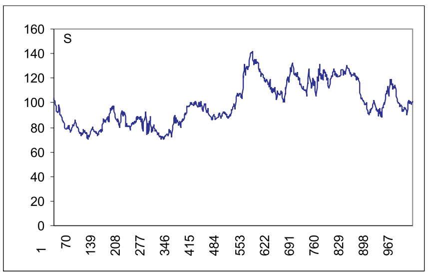
Asset	z	c	PFS	GARCH	Non-rejection region
ABN	0.003	95%	**34**	**19**	$16 < N < 36$
		97.5%	**16**	**13**	$6 < N < 20$
		99%	**7**	**9**	$1 < N < 10$
KPN	0.007	95%	**29**	11	$16 < N < 36$
		97.5%	**14**	**8**	$6 < N < 20$
		99%	**7**	**4**	$1 < N < 10$
JiaLing	0.010	95%	**31**	**22**	$16 < N < 36$
		97.5%	**19**	**14**	$6 < N < 20$
		99%	**8**	**6**	$1 < N < 10$
BaoShan	0.011	95%	**34**	12	$16 < N < 36$
		97.5%	**16**	**8**	$6 < N < 20$
		99%	**7**	**6**	$1 < N < 10$
COSCO	0.004	95%	**32**	14	$16 < N < 36$
		97.5%	**18**	**11**	$6 < N < 20$
		99%	**8**	**5**	$1 < N < 10$
Merchant	0.008	95%	**21**	10	$16 < N < 36$
		97.5%	**12**	5	$6 < N < 20$
		99%	**5**	**4**	$1 < N < 10$

Analyzing financial returns

Erasmus University Rotterdam

Try to determine serially correlated periods





GARCH modeling

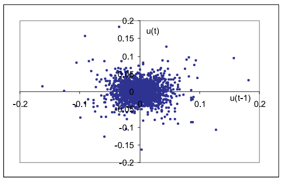
(1) Return $\underline{u}(t) \sim N(\mu, \sigma(t))$

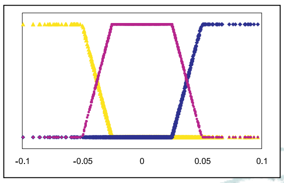
(2) Local volatility is updated conform

$$\underline{\sigma}^2(t) = \gamma \overline{\sigma}^2 + \alpha\, \underline{u}^2(t-1) + \beta \underline{\sigma}^2(t-1)$$

(3) Realistic parameters:

$$\overline{\sigma} = 0.03, \gamma = 0.02, \ \alpha = 0.2, \beta = 0.78$$





Rule base

Future return	very low -0.05	low -0.025	average 0	high 0.025	very high 0.05
Current return					
All	*0.0550*	*0.2265*	*0.4435*	*0.2140*	*0.0610*
Low	0.1271	0.2084	0.2954	0.2302	0.1390
Average	0.0437	0.2293	0.4666	0.2136	0.0468
High	0.1374	0.2077	0.2808	0.2063	0.1679

Rule example:

If *current* return is Low, then a very low or very high *future* return is (very) likely.

Conclusions

- Probabilistic fuzzy systems combine linguistic uncertainty and probabilistic uncertainty
- Very useful in applications where a probabilistic model (pdf estimation) has to be conditioned (or constrained) by linguistic information
- Added value of these models has been demonstrated in various applications

Future research directions

Erasmus University Rotterdam

- New estimation methods for the model parameters
  - Joint estimation
  - Information-theory based techniques
  - Extending maximum likelihood techniques
- Interaction between linguistic knowledge and data-driven estimation
- Optimizing model complexity, model simplification
- Interpretability of probabilistic fuzzy models
- Linguistic descriptions of probability density functions
- New applications

16-Oct-2008 LFA 2008

Bibliography

- J. van den Berg, W. M. van den Bergh, and U. Kaymak. Probabilistic and statistical fuzzy set foundations of competitive exception learning. In *Proceedings of the Tenth IEEE International Conference on Fuzzy Systems*, volume 2, pages 1035–1038, Melbourne, Australia, Dec. 2001.
- J. van den Berg, U. Kaymak, and W.-M. van den Bergh. Probabilistic reasoning in fuzzy rule-based systems. In P. Grzegorzewski, O. Hryniewicz, and M. A. Gil, editors, *Soft Methods in Probability, Statistics and Data Analysis*, Advances in Soft Computing, pages 189–196. Physica Verlag, Heidelberg, 2002.
- U. Kaymak, W.-M. van den Bergh, and J. van den Berg. A fuzzy additive reasoning scheme for probabilistic Mamdani fuzzy systems. In *Proceedings of the 2003 IEEE International Conference on Fuzzy Systems*, volume 1, pages 331–336, St. Louis, USA, May 2003.
- U. Kaymak and J. van den Berg. On probabilistic connections of fuzzy systems. In *Proceedings of the* 15*th Belgium-Netherlands Artificial Intelligence Conference*, pages 187–194, Nijmegen, Netherlands, Oct. 2003.
- J. van den Berg, U. Kaymak, and W.-M. van den Bergh. Financial markets analysis by using a probabilistic fuzzy modelling approach. *International Journal of Approximate Reasoning*, 35 :291–305, 2004.
- L. Waltman, U. Kaymak, and J. van den Berg. Maximum likelihood parameter estimation in probabilistic fuzzy classifiers. In *Proceedings of the 14th Annual IEEE International Conference on Fuzzy Systems*, pages 1098–1103, Reno, Nevada, USA, May 2005.
- D. Xu and U. Kaymak. Value-at-risk estimation by using probabilistic fuzzy systems. In *Proceedings of the 2008 IEEE World Congress on Computational Intelligence (WCCI 2008)*, pages 2109–2116, Hong Kong, June 2008.

Timetable

Submission of full papers:	January 9, 2009
Special session proposals:	December 1, 2008
Acceptance notification:	March 2, 2009
Camera-ready copy due:	April 3, 2009
Early Registration:	April 3, 2009

Conference Secretariat

Center of Intelligent Systems – IDMEC
Instituto Superior Técnico – TU Lisbon
Av. Rovisco Pais, 1049-001 Lisboa, Portugal
E-mail: ifsa-eusflat2009@ist.utl.pt
URL: http://www.ifsa2009.org
URL: http://www.eusflat2009.org