Generalized
Discrete ARMA Models

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

Tobias A. Möller

Christian H. Weiß

Department of Mathematics & Statistics,

Helmut Schmidt University, Hamburg

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

Generalized Discrete ARMA Models

Motivation & Outline

ARMA model for stationary real-valued time series
very popular in theory and applications.
Many attractive features, e. g., ACF $\rho(k) = Corr[X_t, X_{t-k}]$
determined from **Yule-Walker equations**.

Consequently, many attempts to adapt ARMA approach
to time series with range different than $\mathbb{R}$, e. g.,
vector-valued t.s., compositional t.s., integer-valued t.s.,
categorical t.s., . . . (Holan et al., 2010; Weiß, 2018).

Most ARMA-like models tailor-made for particular type of range,
not possible to apply these models to different types of range.

Only universal ARMA-like model:

NDARMA model by Jacobs & Lewis (1983).

For i. i. d. random vectors $\boldsymbol{D}_t = (D_{t,-\mathsf{q}}, \ldots, D_{t,0}, \ldots, D_{t,\mathsf{p}})$ from Mult(1; $\phi_{-\mathsf{q}}, \ldots, \phi_0, \ldots, \phi_{\mathsf{p}}$), one defines

$$X_t \;=\; \sum_{i=1}^{\mathsf{p}} D_{t,i}\, X_{t-i} \;+\; D_{t,0}\, \epsilon_t \;+\; \sum_{j=1}^{\mathsf{q}} D_{t,-j}\, \epsilon_{t-j}.$$

So X_t randomly selects value of one of either $X_{t-1}, \ldots, X_{t-\mathsf{p}}$ or $\epsilon_t, \ldots, \epsilon_{t-\mathsf{q}}$.

NDARMA universally applicable, even to qualitative t.s.
Furthermore, ARMA-like ACF for quantitative range.

But again because of random selection mechanism,
NDARMA sample paths tend to have long runs of values,
interrupted by sudden jumps
$\Rightarrow$ often not appropriate for applications.

More variation in sample paths needed!

In special case of bounded counts, $X_t \in \{0, \ldots, n\}$,
Gouveia et al. (2018) modified NDARMA model
by introducing binomial variation operator (“bvARMA”).

$\Rightarrow$ ...

Idea: To preserve NDARMA's universal applicability and its linear conditional mean,

generalized discrete ARMA (GDARMA) model,

using mean-preserving variation function for additional variation.

Outline:

- Diverse variation functions for different types of t.s. data.
- Unique computation of moments and autocovariances for GDARMA processes.
- Two real applications: integer t.s. of aggregated votes and "CoDa" t.s. concerning television market shares.

HELMUT SCHMIDT
UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH
STAT

Generalized Discrete ARMA Models

Variation functions

Definition of GDARMA(p, q) model:

Let $(\boldsymbol{X}_t)_{\mathbb{Z}}$ be K-dimensional process with state space $\mathcal{S}$.

$(\boldsymbol{\epsilon}_t)_{\mathbb{Z}}$ i. i. d. innovations with $\mathcal{S}$, $\boldsymbol{\epsilon}_t$ independent of $(\boldsymbol{X}_s)_{s<t}$.

$\boldsymbol{D}_t = (D_{t,-\mathsf{q}}, \ldots, D_{t,0}, \ldots, D_{t,\mathsf{p}}) \sim \text{Mult}(1;\ \phi_{-\mathsf{q}}, \ldots, \phi_0, \ldots, \phi_{\mathsf{p}})$,

$\boldsymbol{D}_t$ independent of $(\boldsymbol{\epsilon}_s)_{\mathbb{Z}}$ and of $(\boldsymbol{X}_s)_{s<t}$.

GDARMA(p, q) $(\boldsymbol{X}_t)_{\mathbb{Z}}$ with **variation functions** $\boldsymbol{f}_{\cdot,\cdot} : \mathcal{S} \to \mathcal{S}$ if

$$\boldsymbol{X}_t \ = \ \sum_{i=1}^{\mathsf{p}} D_{t,i}\, \boldsymbol{f}_{t,i}(\boldsymbol{X}_{t-i}) \ + \ D_{t,0}\, \boldsymbol{\epsilon}_t \ + \ \sum_{j=1}^{\mathsf{q}} D_{t,-j}\, \boldsymbol{f}_{t,-j}(\boldsymbol{\epsilon}_{t-j}).$$

GDARMA model very parsimonious,

only $\mathsf{p} + \mathsf{q}$ dependence parameters.

Variation functions:

If $\boldsymbol{f}_{\cdot,\cdot} = \mathrm{id}$ (identity function), then GDARMA = NDARMA.

$\boldsymbol{f}_{\cdot,\cdot}$ assumed to be random,

realized independently of each other and of any other r.v.

To end up with ARMA-like ACF, $\boldsymbol{f}_{\cdot,\cdot}$ **mean-preserving**,

i. e., $E\big[\boldsymbol{f}(\boldsymbol{X}) \mid \boldsymbol{X}\big] = \boldsymbol{X}$ and thus $E\big[\boldsymbol{f}(\boldsymbol{X})\big] = E[\boldsymbol{X}]$.

Then $V\big[f_k(\boldsymbol{X})\big] = V[X_k] + E\big[V[f_k(\boldsymbol{X}) \mid \boldsymbol{X}]\big]$, and

$Cov\big[f_k(\boldsymbol{X}), f_l(\boldsymbol{X})\big] = Cov[X_k, X_l] + E\big[Cov[f_k(\boldsymbol{X}), f_l(\boldsymbol{X}) \mid \boldsymbol{X}]\big]$.

Also $P\big(f(X) \leq k\big) = \sum_x P\big(f(x) \leq k\big)\, P(X = x)$.

Variation functions for counts:

Binomial variation: $f(x) = \mathrm{bv}_n(x)$ for $\mathcal{S} = \{0, \ldots, n\}$,

where $\mathrm{bv}_n(x) \sim \mathrm{Bin}(n, \frac{x}{n})$. We have $V\big[\mathrm{bv}_n(x)\big] = x\left(1 - \frac{x}{n}\right)$.

$\rightarrow$ **bvARMA model** by Gouveia et al. (2018).

Poisson variation: $f(x) = \mathrm{Pv}(x)$ for $\mathcal{S} = \mathbb{N}_0$,

where $\mathrm{Pv}(x) \sim \mathrm{Poi}(x)$. We have $V\big[\mathrm{Pv}(x)\big] = x$.

Geometric variation: $f(x) = \mathrm{gv}(x)$ for $\mathcal{S} = \mathbb{N}_0$,

where $\mathrm{gv}(x) \sim \mathrm{Geom}\left(\frac{1}{1+x}\right)$. We have $V\big[\mathrm{gv}(x)\big] = x\,(1 + x)$.

Variation functions for counts

... with an additional *dispersion parameter*:

Beta-binomial variation: $f(x) = \text{bbv}_{n,\tau}(x)$ for $\mathcal{S} = \{0, \ldots, n\}$,

where $\text{bbv}_{n,\tau}(x) \sim \text{BBin}\big(n; \frac{n-\tau}{\tau-1}\,\frac{x}{n}, \frac{n-\tau}{\tau-1}\,(1-\frac{x}{n})\big)$

with dispersion parameter $\tau \in (0; n)$.

We have $V\big[\text{bbv}_{n,\tau}(x)\big] = \tau\, x\, (1 - \frac{x}{n})$.

Limit $\tau \to 1$ leads to binomial variation.

Negative-binomial variation: $f(x) = \text{nbv}_{\tau}(x)$ for $\mathcal{S} = \mathbb{N}_0$,

where $\text{nbv}_{\tau}(x) \sim \text{NB}\big(\tau, \frac{\tau}{\tau+x}\big)$ with dispersion parameter $\tau > 0$.

We have $V\big[\text{nbv}_{\tau}(x)\big] = \frac{x\,(\tau+x)}{\tau}$.

Variation functions for integers:

Signed bin. var.: $f(x) = \mathrm{sbv}_n(x)$ for $\mathcal{S} = \{-n, \ldots, 0, \ldots, n\}$,

where $\mathrm{sbv}_n(x) \sim \mathrm{sgn}(x)\,\mathrm{Bin}\big(n, \frac{|x|}{n}\big)$; $V\big[\mathrm{sbv}_n(x)\big] = |x|\,\big(1 - \frac{|x|}{n}\big)$.

Signed beta-bin. var.: $f(x) = \mathrm{sbbv}_{n,\tau}(x)$ for $\mathcal{S} = \{-n, \ldots, n\}$,

where $\mathrm{sbbv}_{n,\tau}(x) \sim \mathrm{sgn}(x)\,\mathrm{BBin}\big(n, \frac{n-\tau}{\tau-1}\,\frac{|x|}{n}, \frac{n-\tau}{\tau-1}\,(1 - \frac{|x|}{n})\big)$

with $\tau \in (1; n)$. We have $V\big[\mathrm{sbbv}_{n,\tau}(x)\big] = \tau|x|\,\big(1 - \frac{|x|}{n}\big)$.

Limit $\tau \to 1$ leads to signed binomial variation.

Signed Poisson variation: $f(x) = \mathrm{sPv}(x)$ for $\mathcal{S} = \mathbb{Z}$,

where $\mathrm{sPv}(x) \sim \mathrm{sgn}(x)\,\mathrm{Poi}\big(|x|\big)$. We have $V\big[\mathrm{sPv}(x)\big] = |x|$.

Skellam variation: $(\ldots)$

Variation functions for reals:

Normal variation: $f(x) = \mathsf{nv}_\tau(x)$ for $\mathcal{S} = \mathbb{R}$,

where $\mathsf{nv}_\tau(x) \sim \mathsf{N}(x, \tau)$ with dispersion parameter $\tau > 0$.

We have $V\big[\mathsf{nv}_\tau(x)\big] = \tau$.

Exponential variation: $f(x) = \mathsf{ev}(x)$ for $\mathcal{S} = (0; \infty)$,

where $\mathsf{ev}(x) \sim \mathsf{Exp}(1/x)$ without further parameter.

We have $V\big[\mathsf{ev}(x)\big] = x^2$.

Beta variation: $f(x) = \mathsf{btv}_\tau(x)$ for $\mathcal{S} = (0; 1)$,

where $\mathsf{btv}_\tau(x) \sim \mathsf{Beta}\big(n; \frac{1-\tau}{\tau}\,x, \frac{1-\tau}{\tau}\,(1-x)\big)$ with $\tau > 0$.

We have $V\big[\mathsf{btv}_\tau(x)\big] = \tau x(1-x)$.

Variation functions for categorized data

... if counting units falling into $K+1$ categories:

Multinomial var.: $f(\boldsymbol{x}) = \mathbf{mv}_n(\boldsymbol{x}) \sim \text{Mult}(n; \frac{x_0}{n}, \ldots, \frac{x_K}{n})$.

$\mathbf{mv}_n(\boldsymbol{x})$ generalizes binomial variation.

We have $Cov\Big[\big(\mathbf{mv}_n(\boldsymbol{x})\big)_k, \big(\mathbf{mv}_n(\boldsymbol{x})\big)_l\Big] = \delta_{k,l}\, x_k - \frac{1}{n}\, x_k x_l$.

... if "CoDa" t.s. (Pawlowsky-Glahn & Buccianti, 2011):

Dirichlet variation: $f(\boldsymbol{x}) = \mathbf{Dv}_\tau(\boldsymbol{x})$

for $\mathcal{S} = \big\{\boldsymbol{x} \in (0;1)^{K+1} \mid x_0 + \ldots + x_K = 1\big\}$,

where $\mathbf{Dv}_\tau(\boldsymbol{x}) \sim \text{Dir}(\tau\, x_0, \ldots, \tau\, x_K)$ with disp. par. $\tau > 0$.

We have $Cov\Big[\big(\mathbf{Dv}_\tau(\boldsymbol{x})\big)_k, \big(\mathbf{Dv}_\tau(\boldsymbol{x})\big)_l\Big] = \frac{\delta_{k,l}\, x_k - x_k x_l}{1+\tau}$.

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

Generalized Discrete ARMA Models

Moment properties

Because of mean-preserving property, linear conditional mean

$$E[\boldsymbol{X}_t \mid \boldsymbol{X}_{t-1}, \ldots, \boldsymbol{\epsilon}_{t-1}, \ldots] = \phi_0 \cdot \boldsymbol{\mu}_{\epsilon} + \sum_{i=1}^{p} \phi_i \boldsymbol{X}_{t-i} + \sum_{j=1}^{q} \phi_{-j} \boldsymbol{\epsilon}_{t-j},$$

where $\boldsymbol{\mu}_{\epsilon} := E[\boldsymbol{\epsilon}_t]$. Thus, equality

$$\boldsymbol{\mu} := E[\boldsymbol{X}_t] = E[\boldsymbol{\epsilon}_t] = \boldsymbol{\mu}_{\epsilon}.$$

Yule-Walker-like equations for univariate GDARMA process:

$$\gamma(k) - \sum_{i=1}^{p} \phi_i \gamma(|k-i|) = \sigma_{\epsilon}^2 \sum_{u=k}^{q} \phi_{-u} b_{u-k},$$

where

$$b_k = 0 \text{ for } k < 0, \quad b_0 = \phi_0, \quad b_k = \sum_{i=1}^{p} \phi_i b_{k-i} + \phi_{-k} \text{ for } k > 0.$$

Yule-Walker-like equations for multivariate GDARMA:

$$\boldsymbol{\Gamma}(k) - \sum_{i=1}^{p} \phi_i\, \boldsymbol{\Gamma}(k-i) \;=\; \sum_{u=k}^{q} \phi_{-u}\, \mathbf{A}^{(u-k)},$$

where

$$\mathbf{A}^{(0)} = \phi_0\, \boldsymbol{\Gamma}_{\epsilon}, \qquad \mathbf{A}^{(k)} = \sum_{i=1}^{p} \phi_i\, \mathbf{A}^{(k-i)} + \phi_{-k}\, \boldsymbol{\Gamma}_{\epsilon} \quad \text{for } k > 0.$$

Marginal (co-)variances:

$$V[X] \;=\; \phi_0\, V[\epsilon] \;+\; \phi^{(p)}\, V\big[f(X)\big] \;+\; \phi^{(q)}\, V\big[f(\epsilon)\big],$$

$$\begin{aligned} \gamma_{l,m}(0) \;=\; \phi_0\, \gamma_{\epsilon;\, l,m} &+ \phi^{(p)}\, Cov\Big[\big(\boldsymbol{f}(\boldsymbol{X})\big)_l, \big(\boldsymbol{f}(\boldsymbol{X})\big)_m\Big] \\ &+ \phi^{(q)}\, Cov\Big[\big(\boldsymbol{f}(\boldsymbol{\epsilon})\big)_l, \big(\boldsymbol{f}(\boldsymbol{\epsilon})\big)_m\Big], \end{aligned}$$

where $\quad \phi^{(p)} := \sum_{i=1}^{p} \phi_i, \quad \phi^{(q)} := \sum_{j=1}^{q} \phi_{-j}.$

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH
STAT

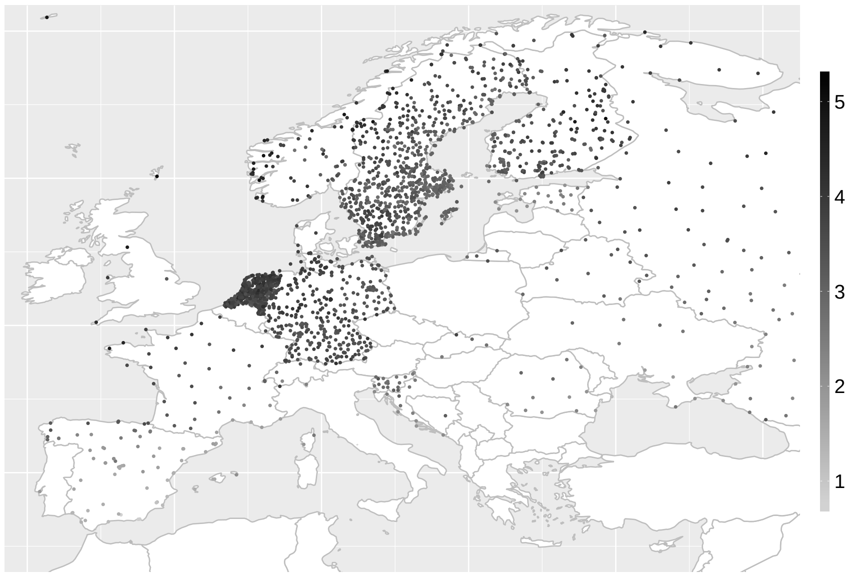
Generalized Discrete ARMA Models

Real applications

Application to bounded counts t.s.:

2565 t.s. with counts of rainy days per week, $\mathcal{S} = \{0, \ldots, 7\}$, at stations in Europe and Russia, see Gouveia et al. (2018).

Used bvARMA(p, q) models with $p, q \leq 1$ plus type of "seasonal" bvARMA model. Outperform NDARMA models.



Means of rainy days t.s.

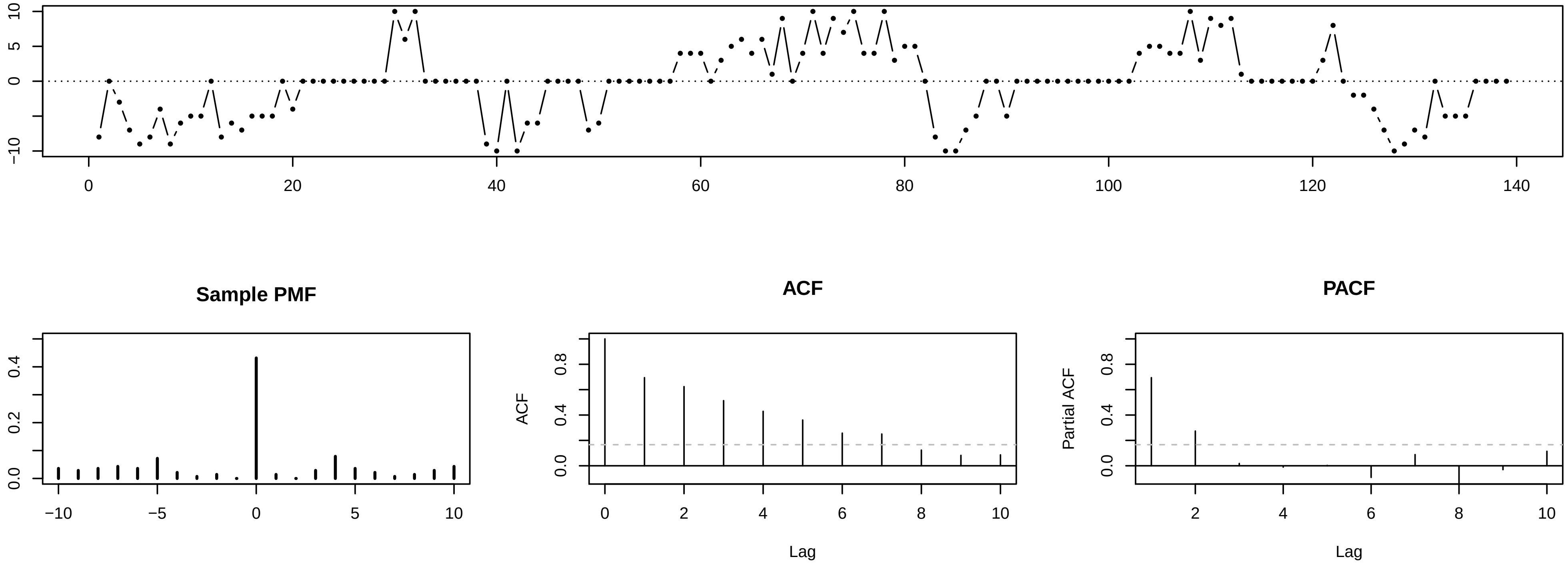
Application to integer t.s.:

Consecutive meetings of Monetary Policy Council (MPC) of the Narodowy Bank Polski (NBP) from 2002 to 2013 (roughly once a month).

$n = 10$ MPC members, may vote for raise of interest rate $(+1)$, or for cut (-1), or for no change (0).

Thus t.s. of length $T = 139$ with range $\{-10, \ldots, 0, \ldots, 10\}$.

Raw data and background information in Sirchenko (2013).



PACF indicates AR-DGP of order ≤ 2.

Sample PMF approx. symmetric, very high zero frequency.

GDARMA candidate models of order $p \leq 2$ and $q \leq 1$,

having signed-ZIB innovations,

either signed binomial or signed beta-binomial variation.

GDARMA(p, q) models using signed binomial variation:

(p, q)	ϕ_{-1}	ϕ_0	ϕ_1	ϕ_2	π	ω		AIC	BIC
(. . .)									
(2, 1)	0.443	0.013	0.342	0.202	0.634	0.422		**584.36**	**599.03**
	(0.078)	-	(0.089)	(0.059)	(0.041)	(0.077)			

GDARMA(p, q) models using signed beta-binomial variation:

(p, q)	ϕ_{-1}	ϕ_0	ϕ_1	ϕ_2	π	ω	τ	AIC	BIC
(. . .)									
(2, 1)	0.230	0.180	0.366	0.224	0.535	0.413	3.076	**577.57**	**595.17**
	(0.119)	-	(0.089)	(0.069)	(0.066)	(0.097)	(0.829)		

Properties of data and fitted GDARMA(p, q) models:

	signed binomial variation				signed beta-binomial variation			
	$E[\lvert X\rvert]$	σ_X^2	$\rho(1)$	$\rho(2)$	$E[\lvert X\rvert]$	σ_X^2	$\rho(1)$	$\rho(2)$
data	3.50	25.11	0.69	0.62	3.50	25.11	0.69	0.62
GDAR(1)	3.64	24.54	0.44	0.20	3.49	25.68	0.53	0.28
GDARMA(1,1)	3.77	27.05	0.43	0.19	3.20	23.44	0.49	0.21
GDAR(2)	3.55	24.70	0.44	0.36	3.26	24.81	0.55	0.47
GDARMA(2,1)	3.67	26.93	0.44	0.35	3.14	23.90	0.51	0.41

Interpretation of preferred GDARMA(2,1) model:

With prob. $\phi_2 = 22.4\,\%$, aggregated votes X_t based on X_{t-2}, with prob. $\phi_1 = 36.6\,\%$ based on X_{t-1}.

New information at time t (innovation ϵ_t) determines voting behavior at same time t with prob. $\phi_0 = 18.0\,\%$, and with delay of one with prob. $\phi_{-1} = 23.0\,\%$.

Application to CoDa t.s.:

Daily shares among $\geq$ 3-year-old viewers in German TV market,

`http://www.quotenmeter.de/c/28/tagesmarktanteile`.

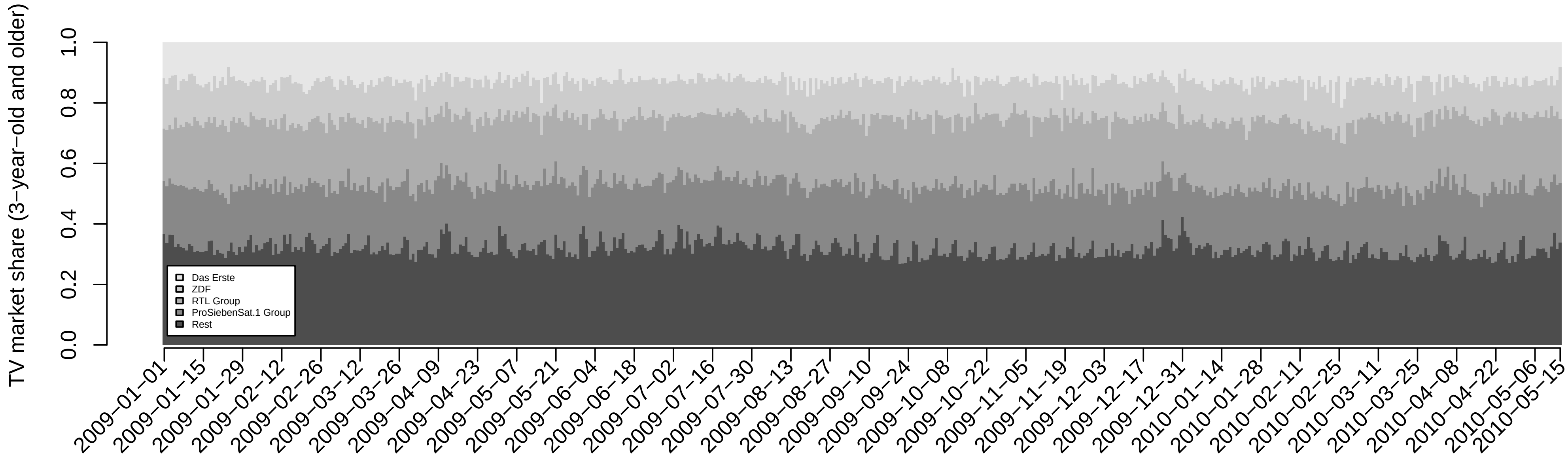
We consider $K = 5$ market participants:

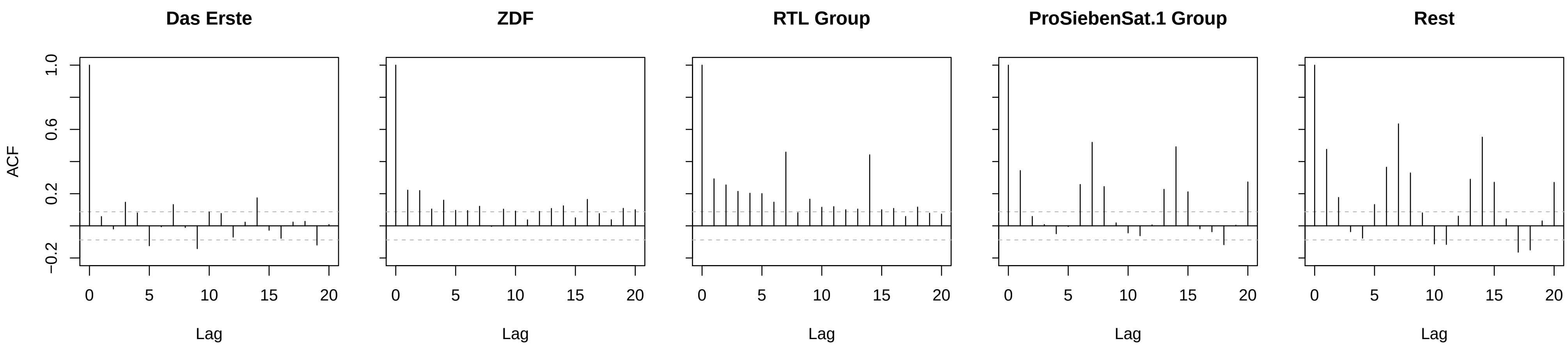
"Das Erste" and "ZDF" (public-sector broadcasts),

"RTL Group", "ProSiebenSat.1 Group" and "Rest",

for period from Jan. 1, 2009 to May 15, 2010 ($T = 500$).

Generalized Discrete ARMA Models





We fit GDAR(p) models of various orders p,

using Dirichlet variation $\mathbf{Dv}_\mathcal{T}(\boldsymbol{x})$ and Dirichlet innovations.

For $\widehat{\boldsymbol{\Gamma}}(k)$ at lags $k = 1, \ldots, 20$,

we calculate $\bar{\rho}(k) = \text{tr}\left(\widehat{\boldsymbol{\Gamma}}(k)\right)/\text{tr}\left(\widehat{\boldsymbol{\Gamma}}(0)\right)$,

as type of weighted mean of individual ACFs.

$\rightarrow$ Serves as overall measure of dependence,

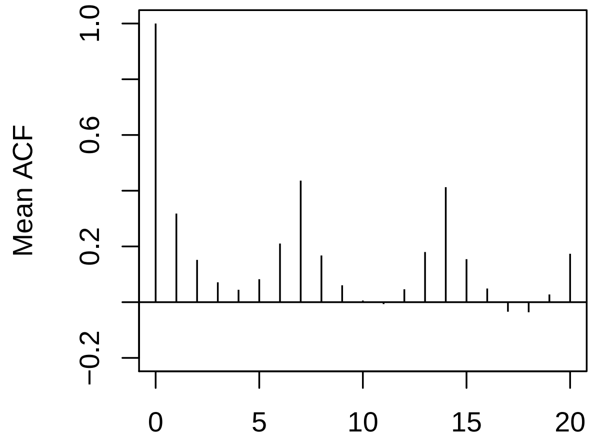
to determine most relevant lags for estimation.

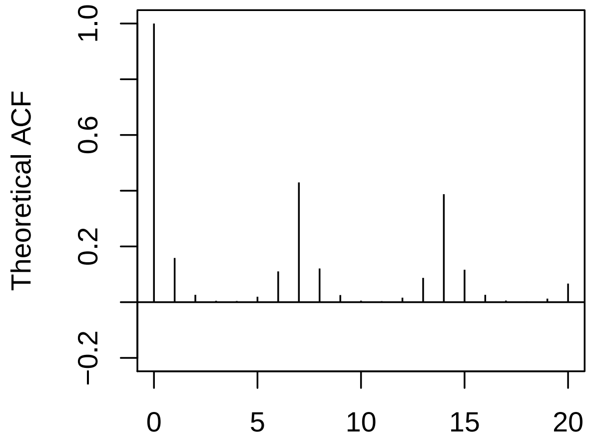
We obtain most relevant lags 7, 14, 1, 6, 13, 20, . . .

Generalized Discrete ARMA Models

Model	ϕ_0	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5	ϕ_6	ϕ_7	ϕ_{14}	α_1	α_2	α_3	α_4	α_5	τ	AIC	BIC
I	0.60							0.40 (0.03)		40.01 (1.79)	38.90 (1.74)	67.71 (3.04)	63.64 (2.87)	99.55 (4.42)	1333.72 (101.85)	−10724	−10694
II	0.43							0.32 (0.03)	0.25 (0.03)	36.84 (2.05)	35.68 (1.99)	61.52 (3.45)	57.59 (3.26)	91.50 (5.07)	1328.10 (99.86)	−10930	−10896
III	0.35	0.10 (0.02)						0.31 (0.03)	0.24 (0.03)	33.28 (2.20)	32.29 (2.12)	55.20 (3.69)	51.95 (3.49)	83.08 (5.46)	1306.33 (91.95)	−10972	−10934
IV	0.32	0.10 (0.02)					0.03 (0.01)	0.31 (0.03)	0.24 (0.03)	32.78 (2.35)	31.68 (2.27)	54.05 (3.94)	50.85 (3.72)	81.74 (5.84)	1261.25 (88.24)	−10978	−10936
V	0.32	0.09 (0.03)	0.02 (0.02)	0.01 (0.01)	0.00 (0.01)	0.00 (0.01)	0.02 (0.02)	0.31 (0.03)	0.23 (0.03)	35.09 (39.35)	34.33 (36.79)	58.48 (70.21)	55.14 (65.75)	88.46 (90.71)	1309.42 (28.39)	−10979	−10920

	μ				σ					$\bar\rho(1)$	$\bar\rho(6)$	$\bar\rho(7)$	$\bar\rho(8)$	$\bar\rho(13)$	$\bar\rho(14)$	$\bar\rho(15)$
	Das Erste	ZDF	RTL	Pro7Sat.1	Das Erste	ZDF	RTL	Pro7Sat.1	Rest							
Data	0.128	0.125	0.221	0.209	0.018	0.018	0.022	0.020	0.028	0.318	0.210	0.436	0.168	0.180	0.413	0.154
I	0.129	0.126	0.219	0.205	0.020	0.020	0.025	0.025	0.028	0.000	0.000	0.398	0.000	0.000	0.158	0.000
II	0.130	0.126	0.217	0.203	0.023	0.022	0.028	0.027	0.031	0.000	0.000	0.421	0.000	0.000	0.383	0.000
III	0.130	0.126	0.216	0.203	0.025	0.024	0.030	0.029	0.034	0.133	0.066	0.416	0.099	0.054	0.377	0.101
IV	0.131	0.126	0.215	0.202	0.025	0.025	0.031	0.030	0.035	0.159	0.110	0.430	0.121	0.087	0.388	0.116
V	0.129	0.126	0.215	0.203	0.024	0.024	0.030	0.029	0.034	0.140	0.095	0.416	0.106	0.075	0.372	0.101





of fitted model IV

Interpretation of fitted model IV:

With prob. $\phi_7 = 31\,\%$, viewers again program from week before, with $\phi_{14} = 24\,\%$ the one from two weeks before.
Program of previous day relevant with $\phi_1 = 10\,\%$, completely new program decision with $\phi_0 = 32\,\%$.
Mean preference for "RTL Group" and "ProSiebenSat.1 Group" twice than for "Das Erste" and "ZDF".

Model IV only uses $4+5+1=10$ model parameters, very parsimonious multivariate time series model.
Certainly, not perfect fit to data, e. g., cannot capture heterogeneity among component-wise ACFs.

Conclusions

- GDARMA offers universal modeling approach
  that maintains linear conditional mean and ARMA-like ACF.
- Applicable to all kinds of data
  by incorporating data-specific variation function.
- GDARMA uses small number of model parameters,
  has unique moment properties.
- Real applications concerning counts, integers and CoDa
  demonstrated wide applicability of GDARMA model.

Thank You

for Your Interest!

Christian H. Weiß

Department of Mathematics & Statistics

Helmut Schmidt University, Hamburg

weissc@hsu-hh.de

Literature

Gouveia et al. (2018) A full ARMA model ... *SERRA* 32, 2495–2514.
Holan et al. (2010) ARMA alphabet soup: ... *Stat. Surveys* 4, 232–274.
Jacobs & Lewis (1983) Stat. discrete autoregressive... *JTSA* 4, 19–36.
Pawlowsky-Glahn & Buccianti (2011) *Comp. Data Analysis: ...* Wiley.
Sirchenko (2013) Modeling status quo decisions: ... *NBP Working p.* 148.
Weiß (2018) *An Introduction to Discrete-Valued Time Series*. Wiley.