Consistent Forecasting Across Time Axes: The B-DARMA as a Time-Shift Operator
Supplementary Material

Harrison Katz, Erica Savage, Kai Thomas Brusch

November 2024

1 A Bayesian Dirichlet Auto-Regressive Moving Average Model (B-DARMA)

1.1 Data Model

We observe a J-component multivariate compositional time series

$$\mathbf{y}_t = (y_{t1}, \dots, y_{tJ})',$$

indexed by $t = 1, \dots, T$, where each component satisfies $0 < y_{tj} < 1$, and the sum of components is one:

$$\mathbf{1}'\mathbf{y}_t = \sum_{j=1}^{J} y_{tj} = 1.$$

We assume a Dirichlet observation model:

$$\mathbf{y}_t \,\Big|\, \boldsymbol{\mu}_t, \phi_t \;\sim\; \text{Dirichlet}\big(\phi_t \, \boldsymbol{\mu}_t\big), \tag{1}$$

where $\boldsymbol{\mu}_t = (\mu_{t1}, \dots, \mu_{tJ})'$ is the mean composition vector with $0 < \mu_{tj} < 1$ and $\mathbf{1}'\boldsymbol{\mu}_t = 1$. The scale parameter is $\phi_t > 0$. Then

$$f(\mathbf{y}_t \mid \boldsymbol{\mu}_t, \phi_t) \;\propto\; \prod_{j=1}^{J} y_{tj}^{\phi_t \, \mu_{tj} - 1}.$$

As $\boldsymbol{\mu}_t$ (and hence $\mathbf{y}_t$) lives in the $(J-1)$-dimensional simplex, we use an *additive log-ratio* (alr) transformation for modeling in an unconstrained space.

1.2 Additive Log-Ratio (alr) Link

Define the alr transform for the mean composition as

$$\boldsymbol{\eta}_t = (\boldsymbol{\mu}_t) \;=\; \left(\log \frac{\mu_{t1}}{\mu_{tj^*}},\; \log \frac{\mu_{t2}}{\mu_{tj^*}},\; \dots,\; \log \frac{\mu_{tJ}}{\mu_{tj^*}}\right), \tag{2}$$

where j^* is a chosen reference component (e.g., $j^* = J$). The dimension of $\boldsymbol{\eta}_t$ is $J - 1$. We invert (2) by

$$\mu_{tj} \;=\; \frac{\exp(\eta_{tj})}{\sum_{\ell=1}^{J} \exp(\eta_{t\ell})}, \quad \text{for } j = 1, \dots, J,$$

setting $\eta_{tj^*} \equiv 0$ in the transform for consistency.

1.3 B-DAR(p, q) Process in alr-Space

We let $\boldsymbol{\eta}_t = (\boldsymbol{\mu}_t)$ be a $(J-1)$-vector that follows a vector auto-regressive moving average process. In the paper, we focus on a B-DAR(1) (or B-DAR(1,1)) for simplicity. Below is the more general B-DAR(p, q) form:

$$\boldsymbol{\eta}_t \;=\; \sum_{p=1}^{P} \mathbf{A}_p \left((\mathbf{y}_{t-p}) - \mathbf{X}_{t-p}\boldsymbol{\beta}\right) \;+\; \sum_{q=1}^{Q} \mathbf{B}_q \left((\mathbf{y}_{t-q}) - \boldsymbol{\eta}_{t-q}\right) \;+\; \mathbf{X}_t\,\boldsymbol{\beta}, \tag{3}$$

for $t = m+1, \ldots, T$, where $m = \max(P, Q)$. Here:

- $\mathbf{A}_p$ and $\mathbf{B}_q$ are $(J-1) \times (J-1)$ coefficient matrices capturing the Vector Auto-Regressive (VAR) and Vector Moving Average (VMA) terms, respectively.
- $\mathbf{X}_t$ is a known $(J-1) \times r_\beta$ matrix of deterministic covariates (including an intercept, seasonality, etc.).
- $\boldsymbol{\beta}$ is an $r_\beta \times 1$ parameter vector.

For a **B-DAR(1)** model (as used in the Airbnb application), we set $P = 1, Q = 0$, yielding

$$\boldsymbol{\eta}_t = \mathbf{A}_1 \left[(\mathbf{y}_{t-1}) - \mathbf{X}_{t-1}\boldsymbol{\beta}\right] \;+\; \mathbf{X}_t\,\boldsymbol{\beta}. \tag{B-DAR(1)}$$

Once $\boldsymbol{\eta}_t$ is obtained, we map back to $\boldsymbol{\mu}_t$ via the inverse-alr transform to get the mean composition.

1.4 Scale Parameter ϕ_t

In addition to $\boldsymbol{\eta}_t$, we model the Dirichlet scale parameter ϕ_t. A log link is typical:

$$\phi_t \;=\; \exp(\mathbf{z}_t\,\boldsymbol{\gamma}), \tag{4}$$

where $\mathbf{z}_t$ is a vector of covariates (possibly the same or a subset of $\mathbf{X}_t$), and $\boldsymbol{\gamma}$ is the associated parameter vector.

1.5 Posterior Inference and Forecasting

Given $\mathbf{y}_{1:T}$, we form a posterior for all unknown parameters

$$\boldsymbol{\theta} = \left(\mathbf{A}_p, \mathbf{B}_q, \boldsymbol{\beta}, \boldsymbol{\gamma}\right),$$

by combining a prior $p(\boldsymbol{\theta})$ with the likelihood implied by (1) and (3). Specifically,

$$p(\boldsymbol{\theta} \mid \mathbf{y}_{1:T}) \;\propto\; p(\boldsymbol{\theta}) \prod_{t=m+1}^{T} p\Big(\mathbf{y}_t \mid \boldsymbol{\eta}_t(\boldsymbol{\theta}), \phi_t(\boldsymbol{\theta})\Big),$$

where $m = \max(P, Q)$. In our Airbnb case study, we fit B-DAR(1) with $\mathbf{X}_t$ capturing seasonality and an intercept, and ϕ_t defined by (4) with a small set of covariates.

To *forecast* $\mathbf{y}_{T+1}, \ldots, \mathbf{y}_{T+S}$, we sample from the posterior of $\boldsymbol{\theta}$ and recursively generate future compositions using the B-DAR(1) structure (or B-DAR(p, q) more generally), then draw $\mathbf{y}_t$ from Dirichlet$(\phi_t\boldsymbol{\mu}_t)$ for $t = T+1, \ldots, T+S$.

2 Airbnb Data Analysis: Model Implementation

2.1 B-DAR(1) Formulation for Lead-Time Distributions

In the Airbnb lead-time forecasting example:

- We let $J = 13$ (i.e., lead-time buckets from 0 to 12 months).
- We set $P = 1$ and $Q = 0$ (so only a single auto-regressive lag).
- $\mathbf{X}_t$ includes an intercept and Fourier terms to capture monthly/annual seasonality.
- $\mathbf{z}_t$ includes an intercept for $\log(\phi_t)$ and the same seasonal terms.

2.2 Priors and MCMC

We place weakly informative normal priors on all elements of $\mathbf{A}_1, \boldsymbol{\beta}, \boldsymbol{\gamma}$. For instance,

$$A_{1,ij} \sim \mathcal{N}(0,1), \quad \beta_k \sim \mathcal{N}(0,1), \quad \gamma_\ell \sim \mathcal{N}(0,1).$$

We then implement HMC sampling in `Stan`. Each iteration yields a draw of $(\mathbf{A}_1, \boldsymbol{\beta}, \boldsymbol{\gamma})$, which we use to compute posterior predictive draws for $\mathbf{y}_t$.

2.3 Forecasting Steps

1. **Fit univariate model for total bookings.** (Part 1 in the paper.)
2. **Transform to compositions.** At each booking date t, the proportion of those bookings that fall into each lead-time bucket is $\mathbf{y}_t$.
3. **Fit B-DAR(1) to these $\mathbf{y}_t$'s.** (Part 2 in the paper.)
4. **Generate out-of-sample predictions.** Use posterior draws from the fitted B-DAR(1) to sample future $\boldsymbol{\eta}_t, \phi_t$ and hence future $\mathbf{y}_t$.
5. **Shift to trip date.** Multiply each day's forecasted total bookings by $\mathbf{y}_t$, summing across booking dates that map to the same trip date or trip month.

References (Selected)

- Katz, H., Brusch, K. T., & Weiss, R. E. (2024). A Bayesian Dirichlet auto-regressive moving average model for forecasting lead times. *International Journal of Forecasting*, 40(4), 1556–1567.
- Zheng, T., & Chen, R. (2017). Dirichlet ARMA models for compositional time series. *Journal of Multivariate Analysis*, 158, 31–46.
- Aitchison, J. (1986). *The Statistical Analysis of Compositional Data.* Chapman & Hall.