Statistical Disaggregation — a Monte Carlo Approach for Imputation under Constraints

Shenggang Hu[1] Hongsheng Dai[2] Fanlin Meng[3] Gareth Roberts[1]

[1]University of Warwick [2] Newcastle University [3] University of Manchester

January 12th, 2024

Imputation/Disaggregation

- Consider a set of random variables $Y^{(1)}, Y^{(2)}, Y^{(3)}, \ldots, Y^{(m)}$ with a known constraint S;

$$S = \sum_{i=1}^{m} Y^{(i)}.$$

Imputation/Disaggregation

- Consider a set of random variables $Y^{(1)}, Y^{(2)}, Y^{(3)}, \ldots, Y^{(m)}$ with a known constraint S;

$$S = \sum_{i=1}^{m} Y^{(i)}.$$

- Problem: impute $Y^{(1)}, \ldots, Y^{(m)}$ given S;

Imputation/Disaggregation

- Consider a set of random variables $Y^{(1)}, Y^{(2)}, Y^{(3)}, \ldots, Y^{(m)}$ with a known constraint S;

$$S = \sum_{i=1}^{m} Y^{(i)}.$$

- Problem: impute $Y^{(1)}, \ldots, Y^{(m)}$ given S;
- Use case: (energy consumption) time series imputation, oversampling;
  - Given a known time series with low resolution (once per day), and a time series with high resolution (48 readings per day) but has missing values (Peppanen et al., 2016);

Imputation/Disaggregation

- Consider a set of random variables $Y^{(1)}, Y^{(2)}, Y^{(3)}, \ldots, Y^{(m)}$ with a known constraint S;

$$S = \sum_{i=1}^{m} Y^{(i)}.$$

- Problem: impute $Y^{(1)}, \ldots, Y^{(m)}$ given S;
- Use case: (energy consumption) time series imputation, oversampling;
  - Given a known time series with low resolution (once per day), and a time series with high resolution (48 readings per day) but has missing values (Peppanen et al., 2016);
  - Given a time series, oversample to estimate a time series with m times the frequency (Allard and Bourotte, 2015);

Constrained Imputation

- Product density given by

$$f_{\mathcal{H}}(\boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(m)}) \propto \left[\prod_{i=1}^{m} f_i(\boldsymbol{y}^{(i)})\right] \mathbb{I}_{\mathcal{H}}(\boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(m)})$$

Constrained Imputation

- Product density given by

$$f_{\mathcal{H}}(\boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(m)}) \propto \left[\prod_{i=1}^{m} f_i(\boldsymbol{y}^{(i)})\right] \mathbb{I}_{\mathcal{H}}(\boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(m)})$$

  - each f_i is easier to sample from;
  - $\mathcal{H}$ is a (linear) equality constraint.

Constrained Imputation

- Product density given by

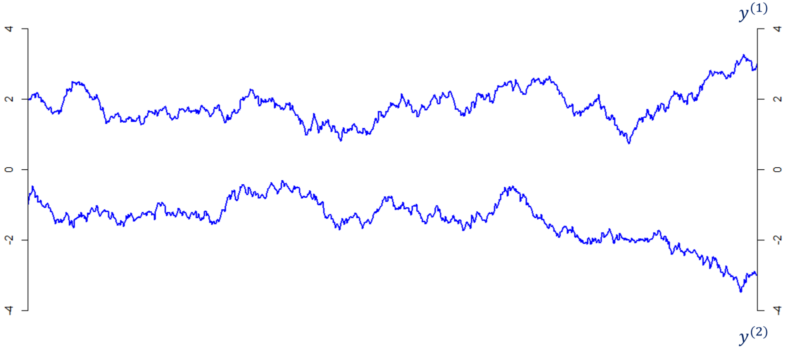
$$f_{\mathcal{H}}(\boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(m)}) \propto \left[\prod_{i=1}^{m} f_i(\boldsymbol{y}^{(i)})\right] \mathbb{I}_{\mathcal{H}}(\boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(m)})$$

  - each f_i is easier to sample from;
  - $\mathcal{H}$ is a (linear) equality constraint.
- Usually intractable;
- Hard to sample directly from an equality constraint.

Key Idea

Two Brownian bridges with restriction on the end points

$$f_{\mathcal{H}}(y^{(1)}, y^{(2)}) \sim \mathcal{N}\left(\boldsymbol{\mu}, \boldsymbol{\Sigma}\right) \delta(y^{(1)} + y^{(2)})$$



$$\text{Brownian bridge} \xrightarrow{\text{Rejection Sampling}} \text{Desired diffusion bridge}$$

What is the Target Distribution/Diffusion?

Biased Diffusion Process

$$
\begin{array}{ccccccc}
 & \text{Start} & & \text{Diffusion bridge} & & \text{End} \\
 & \downarrow & & & & \downarrow \\
\boldsymbol{X}_{s\in[0,T]}^{(i)}: & \boldsymbol{x} & \longrightarrow & X_{s\in[0,T]} | X_0 = \boldsymbol{x}, X_T = \boldsymbol{y} & \longrightarrow & \boldsymbol{y}
\end{array}
$$

Biased Diffusion Process

$$
\begin{array}{cccccccc}
 & & \text{Start} & & \text{Diffusion bridge} & & \text{End} \\
 & & \downarrow & & & & \downarrow \\
\boldsymbol{X}^{(i)}_{s\in[0,T]}: & & \boldsymbol{x} & \longrightarrow & X_{s\in[0,T]} \mid X_0 = \boldsymbol{x}, X_T = \boldsymbol{y} & \longrightarrow & \boldsymbol{y} \\
 & & \uparrow & & & & \uparrow \\
\tilde{\boldsymbol{X}}^{(i)}_{s\in[0,T]}: & & \text{bias} & & & & \text{bias}
\end{array}
$$

Biased Diffusion Process

$$
\begin{tikzcd}
 & \text{Start} \arrow[d] & & \text{Diffusion bridge} & & \text{End} \arrow[d] \\
\boldsymbol{X}^{(i)}_{s\in[0,T]} : & \boldsymbol{x} \arrow[rr] & & X_{s\in[0,T]} | X_0 = \boldsymbol{x}, X_T = \boldsymbol{y} \arrow[rr] & & \boldsymbol{y} \\
\tilde{\boldsymbol{X}}^{(i)}_{s\in[0,T]} : & \text{bias} \arrow[u] & & & & \text{bias} \arrow[u]
\end{tikzcd}
$$

Conditioned on the starting point $\boldsymbol{x}$ and ending point $\boldsymbol{y}$, the diffusion process $\boldsymbol{X}^{(i)}_s$ and the biased process $\tilde{\boldsymbol{X}}^{(i)}_s$ has the same measure.

Biased Diffusion Process

$$
\begin{array}{ccccccc}
 & \text{Start} & & \text{Diffusion bridge} & & \text{End} \\
 & \downarrow & & & & \downarrow \\
\boldsymbol{X}^{(i)}_{s\in[0,T]}: & \boldsymbol{x} & \longrightarrow & X_{s\in[0,T]} | X_0 = \boldsymbol{x}, X_T = \boldsymbol{y} & \longrightarrow & \boldsymbol{y} \\
 & \uparrow & & & & \uparrow \\
\tilde{\boldsymbol{X}}^{(i)}_{s\in[0,T]}: & \text{bias} & & & & \text{bias}
\end{array}
$$

Conditioned on the starting point $\boldsymbol{x}$ and ending point $\boldsymbol{y}$, the diffusion process $\boldsymbol{X}_s^{(i)}$ and the biased process $\tilde{\boldsymbol{X}}_s^{(i)}$ has the same measure.
Set the biased joint distribution on the endpoints to

$$p\left(\tilde{\boldsymbol{X}}_0^{(i)} = \boldsymbol{x}, \tilde{\boldsymbol{X}}_T^{(i)} = \boldsymbol{y}\right) \propto f^2(\boldsymbol{x}) p(\boldsymbol{y}|\boldsymbol{x}) f(\boldsymbol{y})^{-1}.$$

Biased Diffusion Process

$$
\begin{array}{lcccccc}
 & \text{Start} & & \text{Diffusion bridge} & & \text{End} \\
 & \downarrow & & & & \downarrow \\
\boldsymbol{X}^{(i)}_{s\in[0,T]}: & \boldsymbol{x} & \longrightarrow & X_{s\in[0,T]} | X_0 = \boldsymbol{x}, X_T = \boldsymbol{y} & \longrightarrow & \boldsymbol{y} \\
 & \uparrow & & & & \uparrow \\
\tilde{\boldsymbol{X}}^{(i)}_{s\in[0,T]}: & \text{bias} & & & & \text{bias}
\end{array}
$$

Conditioned on the starting point $\boldsymbol{x}$ and ending point $\boldsymbol{y}$, the diffusion process $\boldsymbol{X}_s^{(i)}$ and the biased process $\tilde{\boldsymbol{X}}_s^{(i)}$ has the same measure.
Set the biased joint distribution on the endpoints to

$$p\left(\tilde{\boldsymbol{X}}_0^{(i)} = \boldsymbol{x}, \tilde{\boldsymbol{X}}_T^{(i)} = \boldsymbol{y}\right) \propto f^2(\boldsymbol{x}) p(\boldsymbol{y}|\boldsymbol{x}) f(\boldsymbol{y})^{-1}.$$

Assume $\boldsymbol{X}_s^{(i)}$ is f^2-invariant, the marginal distribution of $\tilde{\boldsymbol{X}}_T^{(i)}$ is

$$p(\tilde{\boldsymbol{X}}_T^{(i)} = \boldsymbol{y}) = f(\boldsymbol{y})$$

Biased Diffusion Process

Target Distribution

Consider an md-dimension diffusion process comprised of m instances of biased diffusion, such that each unbiased process has a different invariant distribution $\propto f_i^2$ and transition density $p_i(\cdot|\cdot)$, $i \in \{1, \ldots, m\}$ respectively.

Biased Diffusion Process

Target Distribution

Consider an md-dimension diffusion process comprised of m instances of biased diffusion, such that each unbiased process has a different invariant distribution $\propto f_i^2$ and transition density $p_i(\cdot|\cdot)$, $i \in \{1, \ldots, m\}$ respectively.
If we impose the equality constraint $\mathcal{H}$ on the endpoints, and require the probability measure induced by the biased diffusion on $(\boldsymbol{X}_0^{(1,\ldots,m)}, \boldsymbol{X}_T^{(1,\ldots,m)}) = (\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(m)}, \boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(m)})$ to follow

$$g_{\mathcal{H}}\left(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(m)}, \boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(m)}\right) \propto \left[\prod_{i=1}^{m} f_i^2(\boldsymbol{x}^{(i)}) p_i(\boldsymbol{y}^{(i)}|\boldsymbol{x}^{(i)}) f_i(\boldsymbol{y}^{(i)})^{-1}\right] \mathbb{I}_{\mathcal{H}}(\boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(m)}).$$

Biased Diffusion Process

Target Distribution

Consider an md-dimension diffusion process comprised of m instances of biased diffusion, such that each unbiased process has a different invariant distribution $\propto f_i^2$ and transition density $p_i(\cdot|\cdot)$, $i \in \{1, \dots, m\}$ respectively.
If we impose the equality constraint $\mathcal{H}$ on the endpoints, and require the probability measure induced by the biased diffusion on $(\boldsymbol{X}_0^{(1,\dots,m)}, \boldsymbol{X}_T^{(1,\dots,m)}) = (\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(m)}, \boldsymbol{y}^{(1)}, \dots, \boldsymbol{y}^{(m)})$ to follow

$$g_{\mathcal{H}}\left(\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(m)}, \boldsymbol{y}^{(1)}, \dots, \boldsymbol{y}^{(m)}\right) \propto \left[\prod_{i=1}^{m} f_i^2(\boldsymbol{x}^{(i)}) p_i(\boldsymbol{y}^{(i)}|\boldsymbol{x}^{(i)}) f_i(\boldsymbol{y}^{(i)})^{-1}\right] \mathbb{I}_{\mathcal{H}}(\boldsymbol{y}^{(1)}, \dots, \boldsymbol{y}^{(m)}).$$

Then the marginal distribution of $\boldsymbol{X}_T = \left(\boldsymbol{X}_T^{(1)}, \dots, \boldsymbol{X}_T^{(m)}\right)$ is exactly the constrained product density

$$f_{\mathcal{H}}(\boldsymbol{y}^{(1)}, \dots, \boldsymbol{y}^{(m)}) \propto \left[\prod_{i=1}^{m} f_i(\boldsymbol{y}^{(i)})\right] \mathbb{I}_{\mathcal{H}}(\boldsymbol{y}^{(1)}, \dots, \boldsymbol{y}^{(m)}).$$

Construction with Langevin Diffusion

Consider a d-dimensional Langevin process $\boldsymbol{X}_s$, $s \in [0, T]$, driven by

$$\mathrm{d}\boldsymbol{X}_s^{(i)} = \nabla \log f_i(\boldsymbol{X}_s)\mathrm{d}s + \mathrm{d}\boldsymbol{W}_s^{(i)},$$

where f_i is the density function of the ith component, and $\boldsymbol{W}_s^{(i)}$ is a d-dimensional Wiener process.

- $\boldsymbol{X}_s^{(i)}$ has invariant distribution $\propto f_i^2(\boldsymbol{x})$

Construction with Langevin Diffusion

Consider a d-dimensional Langevin process $\boldsymbol{X}_s$, $s \in [0, T]$, driven by

$$\mathrm{d}\boldsymbol{X}_s^{(i)} = \nabla \log f_i(\boldsymbol{X}_s)\mathrm{d}s + \mathrm{d}\boldsymbol{W}_s^{(i)},$$

where f_i is the density function of the ith component, and $\boldsymbol{W}_s^{(i)}$ is a d-dimensional Wiener process.

- $\boldsymbol{X}_s^{(i)}$ has invariant distribution $\propto f_i^2(\boldsymbol{x})$

The transition density $p_i(\boldsymbol{x}_T|\boldsymbol{x}_0)$ conditioned on $(\boldsymbol{x}_0, \boldsymbol{x}_T)$ is given by

$$p_i(\boldsymbol{y}|\boldsymbol{x}) \propto \frac{f_i(\boldsymbol{y})}{f_i(\boldsymbol{x})} \times \left(\frac{1}{\sqrt{2\pi T}}\right)^d \exp\left(-\frac{\|\boldsymbol{y}-\boldsymbol{x}\|_2^2}{2T}\right) \mathbb{E}_{\mathbb{W}}\left[\exp\left(-\int_0^T \phi_i(\boldsymbol{x}_s)\mathrm{d}s\right)\right]$$

where $\mathbb{W}$ is the measure induced by the Brownian bridge conditioned on $\boldsymbol{W}_0 = \boldsymbol{x}_0$ and $\boldsymbol{W}_T = \boldsymbol{x}_T$ and

$$\phi_i(\boldsymbol{x}) = \frac{\Delta f_i(\boldsymbol{x})}{2f_i(\boldsymbol{x})}.$$

Now

$$g_{\mathcal{H}}\left(\boldsymbol{x}^{(1)},\dots,\boldsymbol{x}^{(m)},\boldsymbol{y}^{(1)},\dots,\boldsymbol{y}^{(m)}\right) \propto \left[\prod_{i=1}^{m} f_i(\boldsymbol{x}^{(i)})\left(\frac{1}{\sqrt{2\pi T}}\right)^{d}\exp\left(-\frac{\|\boldsymbol{y}^{(i)}-\boldsymbol{x}^{(i)}\|_2^2}{2T}\right)\right] \times \mathbb{E}_{\mathbb{W}}\left[\exp\left(-\int_0^T \phi_i(\boldsymbol{x}_s)\mathrm{d}s\right)\right]\mathbb{I}_{\mathcal{H}}(\boldsymbol{y}^{(1)},\dots,\boldsymbol{y}^{(m)}).$$

Proposal distribution/diffusion?

Proposal and Rejection

Proposal Disribution

Consider the proposal distribution on the end points $(\boldsymbol{X}_0^{(1,\dots,m)}, \boldsymbol{X}_T^{(1,\dots,m)})$ given by

$$h_{\mathcal{H}}\left(\boldsymbol{x}^{(1)},\dots,\boldsymbol{x}^{(m)},\boldsymbol{y}^{(1)},\dots,\boldsymbol{y}^{(m)}\right) \propto \left[\prod_{i=1}^{m} f_i(\boldsymbol{x}^{(i)}) \exp\left(-\frac{||\boldsymbol{y}^{(i)}-\boldsymbol{x}^{(i)}||^2}{2T}\right)\right] \mathbb{I}_{\mathcal{H}}(\boldsymbol{y}^{(1)},\dots,\boldsymbol{y}^{(m)}).$$

Proposal and Rejection

Proposal Disribution

Consider the proposal distribution on the end points $(\boldsymbol{X}_0^{(1,\dots,m)}, \boldsymbol{X}_T^{(1,\dots,m)})$ given by

$$h_{\mathcal{H}}\left(\boldsymbol{x}^{(1)},\dots,\boldsymbol{x}^{(m)},\boldsymbol{y}^{(1)},\dots,\boldsymbol{y}^{(m)}\right) \propto \left[\prod_{i=1}^{m} f_i(\boldsymbol{x}^{(i)}) \exp\left(-\frac{||\boldsymbol{y}^{(i)}-\boldsymbol{x}^{(i)}||^2}{2T}\right)\right] \mathbb{I}_{\mathcal{H}}(\boldsymbol{y}^{(1)},\dots,\boldsymbol{y}^{(m)}).$$

Disregarding the constraint,

$$\frac{g\left(\boldsymbol{x}^{(1)},\dots,\boldsymbol{x}^{(m)},\boldsymbol{y}^{(1)},\dots,\boldsymbol{y}^{(m)}\right)}{h\left(\boldsymbol{x}^{(1)},\dots,\boldsymbol{x}^{(m)},\boldsymbol{y}^{(1)},\dots,\boldsymbol{y}^{(m)}\right)} \propto \mathbb{E}_{\bar{\mathbb{W}}}\left[\exp\left\{-\sum_{i=1}^{m}\int_0^T \phi_i(\boldsymbol{x}_s^{(i)})\mathrm{d}s\right\}\right] \leq 1,$$

where $\bar{\mathbb{W}}$ is the measured induced by a Brownian bridge conditioned on $\boldsymbol{X}_0^{(1:m)} = (\boldsymbol{x}^{(1)},\dots,\boldsymbol{x}^{(m)})$ and $\boldsymbol{X}_T^{(1:m)} = (\boldsymbol{y}^{(1)},\dots,\boldsymbol{y}^{(m)})$.

How to deal with the rejection weight:

$$\mathbb{E}_{\bar{\mathbb{W}}}\left[\exp\left\{-\sum_{i=1}^{m}\int_0^T \phi_i(\boldsymbol{x}_s^{(i)})\mathrm{d}s\right\}\right]$$

One-Sample Rejection Step

Poisson Point Process

Suppose that $0 < \phi < M$, and let Φ be a Poisson point process of intensity 1 defined on the space $[0, T] \times [0, M]$.
Let $A := \{(t, u) \in [0, T] \times [0, M] : u \leq \phi(\omega_t)\}$ denote the region under curve $\phi(\omega_t)$, $\omega_t \sim \bar{\mathbb{W}}$, then

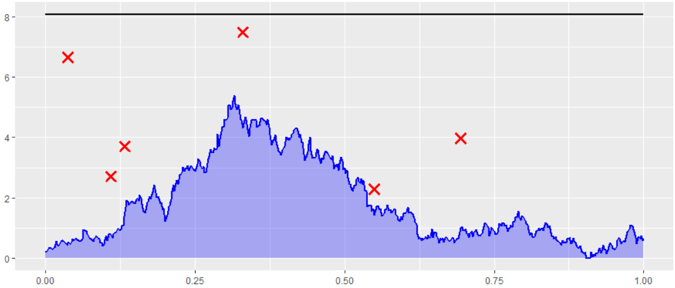
$$\mathbb{P}(N_\Phi(A) = 0|\omega) = \exp\left\{-\int_0^T \phi(\omega_s)\mathrm{d}s\right\}$$

One-Sample Rejection Step

Poisson Point Process

Suppose that $0 < \phi < M$, and let Φ be a Poisson point process of intensity 1 defined on the space $[0, T] \times [0, M]$.
Let $A := \{(t, u) \in [0, T] \times [0, M] : u \leq \phi(\omega_t)\}$ denote the region under curve $\phi(\omega_t)$, $\omega_t \sim \bar{\mathbb{W}}$, then

$$\mathbb{P}(N_\Phi(A) = 0|\omega) = \exp\left\{-\int_0^T \phi(\omega_s)\mathrm{d}s\right\}$$



Is $\phi = \sum_i \frac{\Delta f_i}{2 f_i}$ bounded?

Layered Brownian Bridge

Deciding on the layer also determines the bounds on $X_t, t \in [0, T]$ and subsequently $\phi(X_t)$.

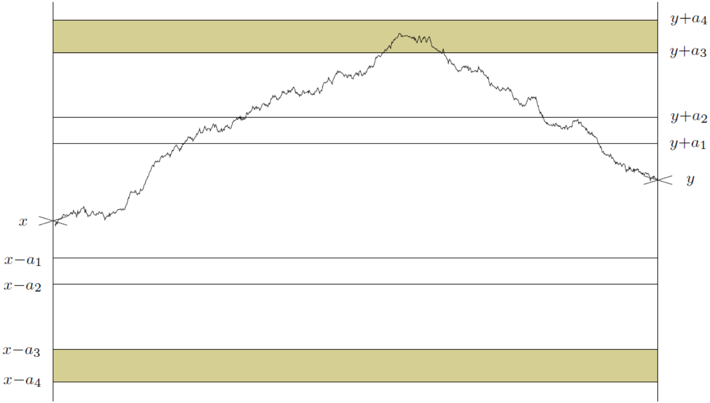


Figure: A sample path from $X_0 = x$ to $X_T = y$. The trajectory landed in the fourth layer. (Beskos et al., 2008)

Are the following definitions mathematically rigorous?

$$f_{\mathcal{H}}(\boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(m)}) \propto \left[\prod_{i=1}^{m} f_i(\boldsymbol{y}^{(i)})\right] \mathbb{I}_{\mathcal{H}}(\boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(m)})$$

$$g_{\mathcal{H}}\left(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(m)}, \boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(m)}\right) \propto \left[\prod_{i=1}^{m} f_i^2(\boldsymbol{x}^{(i)}) p_i(\boldsymbol{y}^{(i)} | \boldsymbol{x}^{(i)}) f_i(\boldsymbol{y}^{(i)})^{-1}\right] \mathbb{I}_{\mathcal{H}}(\boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(m)})$$

$$h_{\mathcal{H}}\left(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(m)}, \boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(m)}\right) \propto \left[\prod_{i=1}^{m} f_i(\boldsymbol{x}^{(i)}) \exp\left(-\frac{||\boldsymbol{y}^{(i)} - \boldsymbol{x}^{(i)}||^2}{2T}\right)\right] \mathbb{I}_{\mathcal{H}}(\boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(m)})$$

Existence of Equality-Constrained Density

Lemma (Regular Value Theorem)

Let $\vec{\boldsymbol{h}} : \mathbb{R}^{n+k} \to \mathbb{R}^k$, $0 < k < md$ *be a smooth function such that* $\forall \boldsymbol{u} \in \vec{\boldsymbol{h}}^{-1}(0)$, *the derivative* $\mathtt{d}\vec{\boldsymbol{h}}_{\boldsymbol{u}} : \mathbb{R}^{n+k} \to \mathbb{R}^k$ *is surjective. Then, the set*

$$\mathcal{H} := \vec{\boldsymbol{h}}^{-1}(\mathbf{0}) = \left\{ \boldsymbol{u} \in \mathbb{R}^{n+k} : \vec{\boldsymbol{h}}(\boldsymbol{u}) = \mathbf{0} \right\},$$

is a n-dimensional manifold. Moreover, there exists a **canonical** *volume form* $Vol_{\mathcal{H}}$ *defined on* $\mathcal{H}$ *such that*

$$\int_{\mathcal{H}} \mathtt{d}Vol_{\mathcal{H}} = \textit{ Volume of } \mathcal{H}.$$

Existence of Equality-Constrained Density

Theorem

If $\mathcal{H} \subset \mathbb{R}^{n+k}$ is an n-dimensional manifold, then the following holds:

1. *we may define naturally the measures $P_f, P_g : \mathcal{B}(\mathcal{H}) \to [0,1]$ induced by restricting f, g on $\mathcal{H}$ such that the Radon-Nikodym derivative with respect to the volume measure $\int_\cdot \mathrm{d}Vol_{\mathcal{H}}$ is proportional to their corresponding density on the full space;*
2. *if $f \ll g$, then $P_f \ll P_g$ with*

$$\frac{\mathrm{d}P_f}{\mathrm{d}P_g} \propto \frac{f}{g}.$$

Thus indeed,

$$\frac{g_{\mathcal{H}}\left(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(m)}, \boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(m)}\right)}{h_{\mathcal{H}}\left(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(m)}, \boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(m)}\right)} \propto \mathbb{E}_{\bar{\mathbb{W}}}\left[\exp\left\{-\sum_{i=1}^{m} \int_0^T \phi_i(\boldsymbol{x}_s^{(i)}) \mathrm{d}s\right\}\right] \mathbb{I}_{\mathcal{H}}(\boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(m)}).$$

Sampling from Proposal

Recall the proposal distribution

$$h_{\mathcal{H}}\left(\boldsymbol{x}^{(1)},\ldots,\boldsymbol{x}^{(m)},\boldsymbol{y}^{(1)},\ldots,\boldsymbol{y}^{(m)}\right) \propto \left[\prod_{i=1}^{m} f_i(\boldsymbol{x}^{(i)}) \exp\left(-\frac{||\boldsymbol{y}^{(i)}-\boldsymbol{x}^{(i)}||^2}{2T}\right)\right] \mathbb{I}_{\mathcal{H}}(\boldsymbol{y}^{(1)},\ldots,\boldsymbol{y}^{(m)}).$$

- If $\mathcal{N}(\boldsymbol{x}, TI_{md})\mathbb{I}_{\mathcal{H}}(\boldsymbol{y}^{(1)},\ldots,\boldsymbol{y}^{(m)})$ can be analytically simplified, e.g.,
  - linear constraint $\rightarrow$ Gaussian distribution;
  - Spherical/Elliptical constraint $\rightarrow$ von Mises-Fisher distribution.

Sampling from Proposal

Recall the proposal distribution

$$h_{\mathcal{H}}\left(\boldsymbol{x}^{(1)},\ldots,\boldsymbol{x}^{(m)},\boldsymbol{y}^{(1)},\ldots,\boldsymbol{y}^{(m)}\right) \propto \left[\prod_{i=1}^{m} f_i(\boldsymbol{x}^{(i)}) \exp\left(-\frac{||\boldsymbol{y}^{(i)}-\boldsymbol{x}^{(i)}||^2}{2T}\right)\right] \mathbb{I}_{\mathcal{H}}(\boldsymbol{y}^{(1)},\ldots,\boldsymbol{y}^{(m)}).$$

- If $\mathcal{N}(\boldsymbol{x}, TI_{md})\mathbb{I}_{\mathcal{H}}(\boldsymbol{y}^{(1)},\ldots,\boldsymbol{y}^{(m)})$ can be analytically simplified, e.g.,
  - linear constraint $\rightarrow$ Gaussian distribution;
  - Spherical/Elliptical constraint $\rightarrow$ von Mises-Fisher distribution.

  Sample $\boldsymbol{y} \sim \mathcal{N}(\boldsymbol{x}, TI_{md})\mathbb{I}_{\mathcal{H}}(\boldsymbol{y}^{(1)},\ldots,\boldsymbol{y}^{(m)})$ and correct for $\int_{\mathcal{H}} \mathcal{N}(\boldsymbol{x}, TI_{md})\mathbb{I}_{\mathcal{H}}(\boldsymbol{y}^{(1)},\ldots,\boldsymbol{y}^{(m)})\mathrm{dVol}_{\mathcal{H}}$.

Sampling from Proposal

Recall the proposal distribution

$$h_{\mathcal{H}}\left(\boldsymbol{x}^{(1)},\ldots,\boldsymbol{x}^{(m)},\boldsymbol{y}^{(1)},\ldots,\boldsymbol{y}^{(m)}\right)\propto\left[\prod_{i=1}^{m}f_i(\boldsymbol{x}^{(i)})\exp\left(-\frac{\|\boldsymbol{y}^{(i)}-\boldsymbol{x}^{(i)}\|^2}{2T}\right)\right]\mathbb{I}_{\mathcal{H}}(\boldsymbol{y}^{(1)},\ldots,\boldsymbol{y}^{(m)}).$$

- If $\mathcal{N}(\boldsymbol{x}, TI_{md})\mathbb{I}_{\mathcal{H}}(\boldsymbol{y}^{(1)},\ldots,\boldsymbol{y}^{(m)})$ can be analytically simplified, e.g.,
  - linear constraint $\to$ Gaussian distribution;
  - Spherical/Elliptical constraint $\to$ von Mises-Fisher distribution.

  Sample $\boldsymbol{y}\sim\mathcal{N}(\boldsymbol{x}, TI_{md})\mathbb{I}_{\mathcal{H}}(\boldsymbol{y}^{(1)},\ldots,\boldsymbol{y}^{(m)})$ and correct for $\int_{\mathcal{H}}\mathcal{N}(\boldsymbol{x}, TI_{md})\mathbb{I}_{\mathcal{H}}(\boldsymbol{y}^{(1)},\ldots,\boldsymbol{y}^{(m)})\mathrm{dVol}_{\mathcal{H}}$.

- For arbitrary manifold constraint,
  1. Sample $\boldsymbol{x}^{(i)}\sim f_i$;
  2. Sample $\boldsymbol{y}^{(1:m)}$ uniformly from $\mathcal{H}$ (e.g., by constrained HMC);
  3. Correct for $\exp(-\|\boldsymbol{y}^{(i)}-\boldsymbol{x}^{(i)}\|^2/2T)$;

Daily to Tridaily Imputation

- Coming back to the imputation problem on energy consumption:

$$\begin{array}{ccccccc} \text{total} & & \text{peak} & & \text{off-peak} & & \text{night} \\ \downarrow & & \downarrow & & \downarrow & & \downarrow \\ S_t & = & Y_t^{(1)} & + & Y_t^{(2)} & + & Y_t^{(3)} \end{array}$$

  where S_t is the total consumption on the day t.
- The time series S_t is known,
- impute the segmented consumption $Y_t^{(i)}$ given S_t.

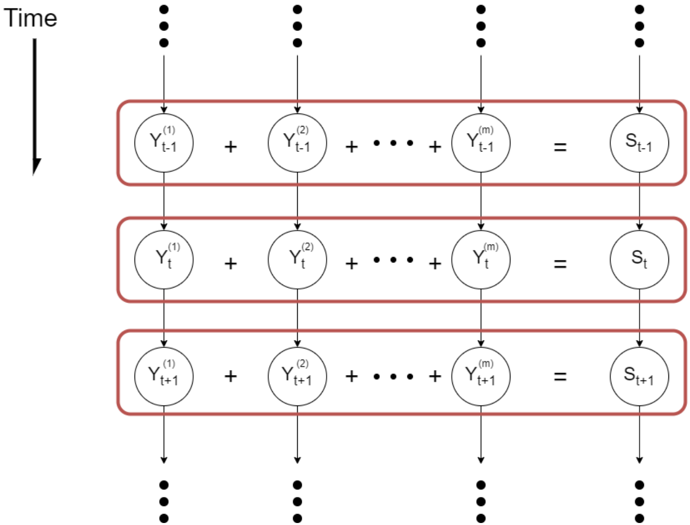
Daily to Tridaily Imputation

- Coming back to the imputation problem on energy consumption:

$$\begin{array}{ccccccc} \text{total} & & \text{peak} & & \text{off-peak} & & \text{night} \\ \downarrow & & \downarrow & & \downarrow & & \downarrow \\ S_t & = & Y_t^{(1)} & + & Y_t^{(2)} & + & Y_t^{(3)} \end{array}$$

  where S_t is the total consumption on the day t.
- The time series S_t is known,
- impute the segmented consumption $Y_t^{(i)}$ given S_t.
- Implemented model: treat $Y_t^{(1)}$, $Y_t^{(2)}$, $Y_t^{(3)}$ as separate time series.

Daily to Tridaily Imputation



Model

Definition (Generalized Logistic Distribution)

Let $\alpha, \beta, \gamma > 0$, $C \in \mathbb{R}$, and $X_1 \sim \Gamma(\alpha, 1)$, $X_2 \sim \Gamma(\beta, 1)$. Let

$$Y := \gamma \log \left(\frac{X_1}{X_2} \right) + C,$$

then X is said to follow a Generalized Logistic distribution with parameter $(\alpha, \beta, \gamma, C)$, denoted $X \sim \text{GenLog}(\alpha, \beta, \gamma, C)$.

Model

Definition (Generalized Logistic Distribution)

Let $\alpha, \beta, \gamma > 0$, $C \in \mathbb{R}$, and $X_1 \sim \Gamma(\alpha, 1)$, $X_2 \sim \Gamma(\beta, 1)$. Let

$$Y := \gamma \log\left(\frac{X_1}{X_2}\right) + C,$$

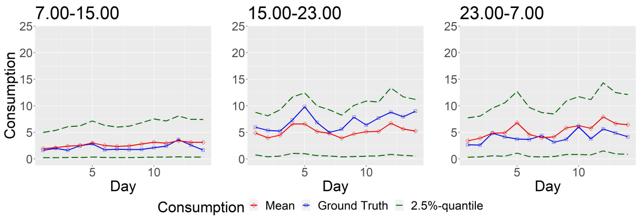
then X is said to follow a Generalized Logistic distribution with parameter $(\alpha, \beta, \gamma, C)$, denoted $X \sim \mathrm{GenLog}(\alpha, \beta, \gamma, C)$.

Autoregressive model for $Y_t^{(i)}$ where Ξ_t are the extra regressors,

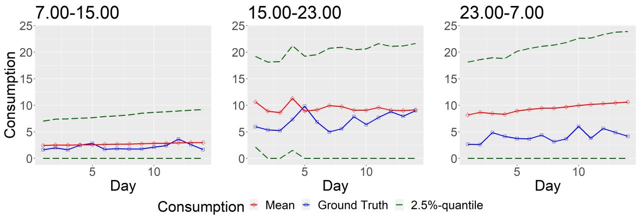
$$Y_t^{(i)} \sim \mathrm{GenLog}\left(\alpha^{(j)}, \beta^{(j)}, \gamma^{(j)}, C^{(j)} + \mu_t^{(j)}\right), \qquad \mu_t^{(j)} = \sum_{r=1}^{K} \Phi_r^{(j)} Y_{t-r}^{(j)} + \Xi_t \psi^{(j)}$$

such that $\mathbb{E}[Y_t^{(i)} | Y_{s<t}^{(i)}] = \mu_t^{(i)}$.

Daily to Tridaily Imputation



(a) Constrained simulation



(b) Unconstrained simulation

Peak-Trough Estimation

- Predict the peak and trough consumption of at 6-minute level in each 30-minute intervals:

$$\underset{\substack{\text{30-min total}\\ \downarrow}}{S_t} = \underset{\substack{\text{interval 1}\\ \downarrow}}{Y_t^{(1)}} + \underset{\substack{\text{interval 2}\\ \downarrow}}{Y_t^{(2)}} + \underset{\substack{\text{interval 3}\\ \downarrow}}{Y_t^{(3)}} + \underset{\substack{\text{interval 4}\\ \downarrow}}{Y_t^{(4)}} + \underset{\substack{\text{interval 5}\\ \downarrow}}{Y_t^{(5)}}$$

where each interval represents a reading average of 6 minutes.

Peak-Trough Estimation

- Predict the peak and trough consumption of at 6-minute level in each 30-minute intervals:

$$\underbrace{S_t}_{\text{30-min total}} = \underbrace{Y_t^{(1)}}_{\text{interval 1}} + \underbrace{Y_t^{(2)}}_{\text{interval 2}} + \underbrace{Y_t^{(3)}}_{\text{interval 3}} + \underbrace{Y_t^{(4)}}_{\text{interval 4}} + \underbrace{Y_t^{(5)}}_{\text{interval 5}}$$

  where each interval represents a reading average of 6 minutes.
- Want to predict $\min\{Y_t^{(1:5)}\}$ and $\max\{Y_t^{(1:5)}\}$ for every 30-minute period given S_t.
- Use the average reading S_t as the baseline.

Peak-Trough Estimation

References

D. Allard and M. Bourotte. Disaggregating Daily Precipitations into Hourly Values with a Transformed Censored Latent Gaussian Process. *Stochastic Environmental Research and Risk Assessment*, 29(2): 453–462, 2015.

A. Beskos, O. Papaspiliopoulos, and G. O. Roberts. A Factorisation of Diffusion Measure and Finite Sample Path Constructions. *Methodology and Computing in Applied Probability*, 10(1):85–104, 2008.

J. Peppanen, X. Zhang, S. Grijalva, and R. Matthew. Handling Bad or Missing Smart Meter Data through Advanced Data Imputation. *Conference: IEEE PES Innovative Smart Grid Technologies (ISGT), September, DOI: 10.1109/ISGT.2016.7781213*, 2016.

Consider sampling three shifted T-distribution $X_1 \sim T_{2.01}(-2)$, $X_2 \sim T_{2.01}(3)$ and $X_3 \sim T_{2.01}(5)$ subject to the constraint $X_1 + X_2 + X_3 = 10$.