DEPARTMENT *of* BIOSTATISTICS EPIDEMIOLOGY & INFORMATICS

Bayesian Functional Mixed Model using Non-Linear Representations (Bayes-FMM-NL)

Edward Gunning, Giles Hooker and Jeffrey Morris

WIT meeting

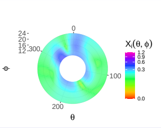
September 2025

Functional Mixed Models (FMMs)

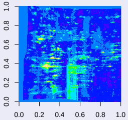
Functional Data

Figures/gait-data-plot.pdf

(a)



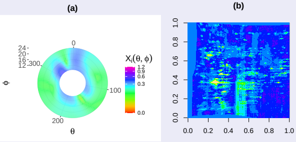
(b)



Functional Mixed Models (FMMs)

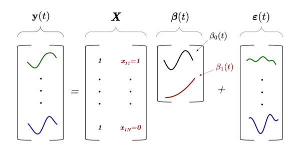
Functional Data

Figures/gait-data-plo



Regression Outcomes

Regress the objects on scalar (or functional) covariates.

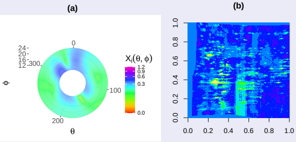


$\rightarrow$ inference, prediction (and data generation!).

Functional Mixed Models (FMMs)

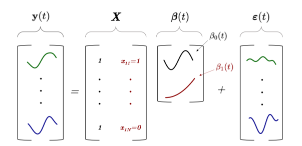
Functional Data

Figures/gait-data-plo



Regression Outcomes

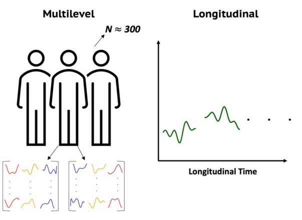
Regress the objects on scalar (or functional) covariates.



$\rightarrow$ inference, prediction (and data generation!).

Dependence Structures

Account for multilevel, longitudinal, spatial, etc. dependence.



Functional Mixed Model

Functional Mixed Model:

$$\mathbf{Y}(t) = \mathbf{X}\mathbf{B}(t) + \mathbf{Z}\mathbf{U}(t) + \mathbf{E}(t)$$

often measured on grid $\mathbf{t} = (0 = t_1, t_2, \ldots, t_D = T)$.

Functional Mixed Model

Functional Mixed Model:

$$\mathbf{Y}(t) = \mathbf{X}\mathbf{B}(t) + \mathbf{Z}\mathbf{U}(t) + \mathbf{E}(t)$$

often measured on grid $\mathbf{t} = (0 = t_1, t_2, \ldots, t_D = T)$.

- **Cui et al. (2022)**: Fit pointwise models

$$\mathbf{Y}_{\cdot,t_j} = \mathbf{X}\mathbf{B}_{\cdot,t_j} + \mathbf{Z}\mathbf{U}_{\cdot,t_j} + \mathbf{E}_{\cdot,t_j}, \quad j = 1, \ldots, D,$$

  and post-smooth the pointwise estimates to construct smooth $\widehat{\mathbf{B}}(t)$.

Functional Mixed Model

Functional Mixed Model:

$$\mathbf{Y}(t) = \mathbf{XB}(t) + \mathbf{ZU}(t) + \mathbf{E}(t)$$

often measured on grid $\mathbf{t} = (0 = t_1, t_2, \ldots, t_D = T)$.

- **Cui et al. (2022)**: Fit pointwise models

  $$\mathbf{Y}_{\cdot,t_j} = \mathbf{XB}_{\cdot,t_j} + \mathbf{ZU}_{\cdot,t_j} + \mathbf{E}_{\cdot,t_j}, \quad j = 1, \ldots, D,$$

  and post-smooth the pointwise estimates to construct smooth $\widehat{\mathbf{B}}(t)$.
- **Scheipl, Staicu, and Greven (2015) and Guo (2002)**: Basis approximations $\boldsymbol{\beta}_a(t) \approx \sum_k \beta^*_{ak}\phi^B_k(t)$ and $\mathbf{u}_i(t) \approx \sum_k u^*_{ik}\phi^U_k(t)$, so that:

  $$\text{vec}(\mathbf{Y}(\mathbf{t})) = \boldsymbol{\Phi}^X\boldsymbol{\beta}^* + \boldsymbol{\Phi}^U\boldsymbol{u}^* + \boldsymbol{\varepsilon},$$

  as a standard penalized scalar LMM.

Functional Mixed Model

Functional Mixed Model:

$$\mathbf{Y}(t) = \mathbf{X}\mathbf{B}(t) + \mathbf{Z}\mathbf{U}(t) + \mathbf{E}(t)$$

often measured on grid $\mathbf{t} = (0 = t_1, t_2, \ldots, t_D = T)$.

- **Cui et al. (2022)**: Fit pointwise models

$$\mathbf{Y}_{\cdot,t_j} = \mathbf{X}\mathbf{B}_{\cdot,t_j} + \mathbf{Z}\mathbf{U}_{\cdot,t_j} + \mathbf{E}_{\cdot,t_j}, \quad j = 1, \ldots, D,$$

  and post-smooth the pointwise estimates to construct smooth $\widehat{\mathbf{B}}(t)$.
- **Scheipl, Staicu, and Greven (2015) and Guo (2002)**: Basis approximations $\boldsymbol{\beta}_a(t) \approx \sum_k \beta^*_{ak}\phi^B_k(t)$ and $\mathbf{u}_i(t) \approx \sum_k u^*_{ik}\phi^U_k(t)$, so that:

$$\text{vec}(\mathbf{Y}(\mathbf{t})) = \boldsymbol{\Phi}^X\boldsymbol{\beta}^* + \boldsymbol{\Phi}^U\boldsymbol{u}^* + \boldsymbol{\varepsilon},$$

  as a standard penalized scalar LMM.
- **Morris and Carroll (2006)**: Near-lossless basis representation for the $\mathbf{Y}(t) \approx \mathbf{Y}^*\boldsymbol{\Phi}(t)$ and all other terms, then model the basis coefficients $\mathbf{Y}^*$ using a multivariate Bayesian mixed model.

Bayesian Functional Mixed Model

BayesFMM modeling framework (Morris and Carroll, 2006; Morris, Baladandayuthapani, et al., 2011; Morris, 2017):

$\mathbf{Y}(\mathbf{t})$

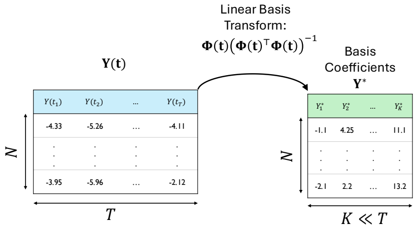
$Y(t_1)$	$Y(t_2)$	...	$Y(t_T)$
-4.33	-5.26	...	-4.11
⋮	⋮	⋮	⋮
-3.95	-5.96	...	-2.12

N

T

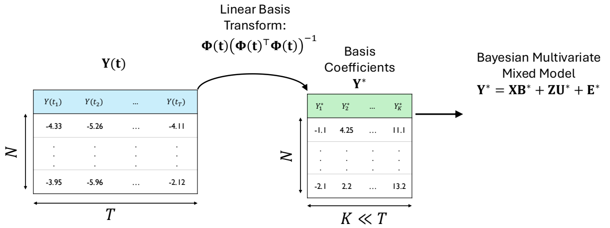
Bayesian Functional Mixed Model

BayesFMM modeling framework (Morris and Carroll, 2006; Morris, Baladandayuthapani, et al., 2011; Morris, 2017):



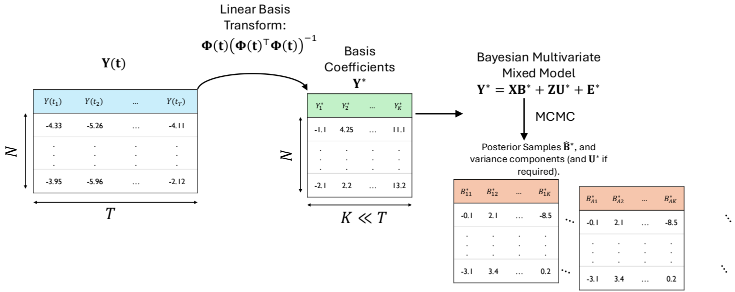
Bayesian Functional Mixed Model

`BayesFMM` modeling framework (Morris and Carroll, 2006; Morris, Baladandayuthapani, et al., 2011; Morris, 2017):



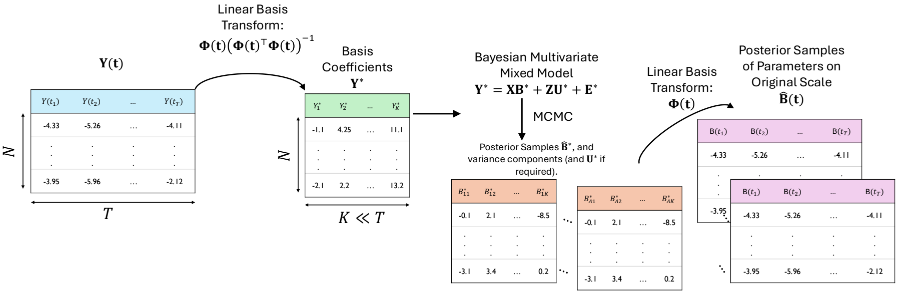
Bayesian Functional Mixed Model

BayesFMM modeling framework (Morris and Carroll, 2006; Morris, Baladandayuthapani, et al., 2011; Morris, 2017):



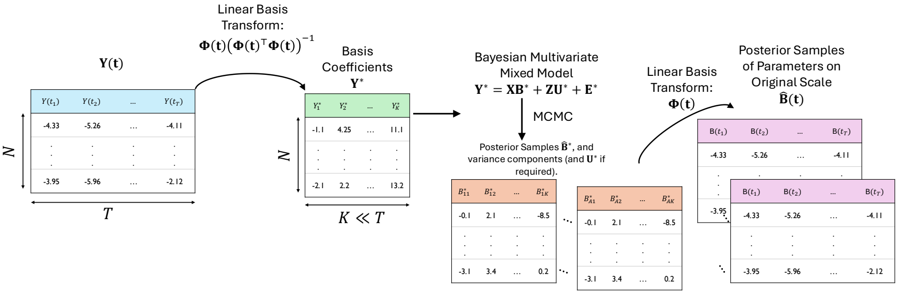
Bayesian Functional Mixed Model

BayesFMM modeling framework (Morris and Carroll, 2006; Morris, Baladandayuthapani, et al., 2011; Morris, 2017):



Bayesian Functional Mixed Model

BayesFMM modeling framework (Morris and Carroll, 2006; Morris, Baladandayuthapani, et al., 2011; Morris, 2017):



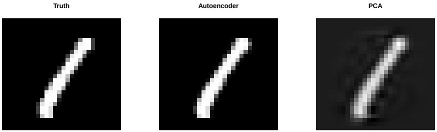
with different bases $\{\phi_K(t)\}_{k=1}^{K}$ for different data types. E.g., wavelets, FPCs.

Bayesian Functional Mixed Model

Linear basis representations may not provide the most compact or efficient representation (see, e.g., Zohner et al., 2025):

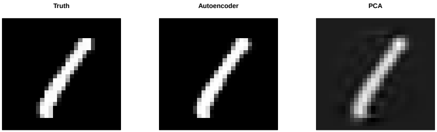
Bayesian Functional Mixed Model

Linear basis representations may not provide the most compact or efficient representation (see, e.g., Zohner et al., 2025):



Bayesian Functional Mixed Model

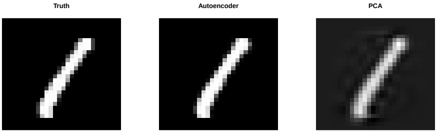
Linear basis representations may not provide the most compact or efficient representation (see, e.g., Zohner et al., 2025):



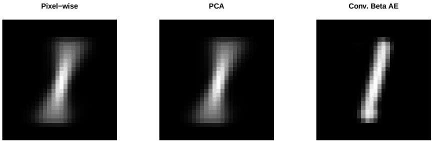
And their transformed spaces might be more amenable to statistical modeling:

Bayesian Functional Mixed Model

Linear basis representations may not provide the most compact or efficient representation (see, e.g., Zohner et al., 2025):



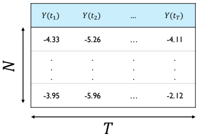
And their transformed spaces might be more amenable to statistical modeling:



Bayes-FMM-NL

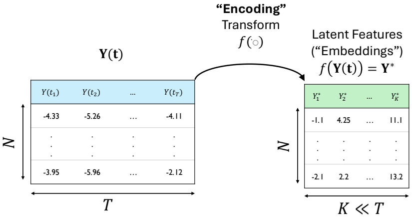
Embed general, potentially **non-linear** encoding $f(\cdot)$ and decoding $g(\cdot)$ transformations into the `BayesFMM` framework:

$\mathbf{Y}(\mathbf{t})$



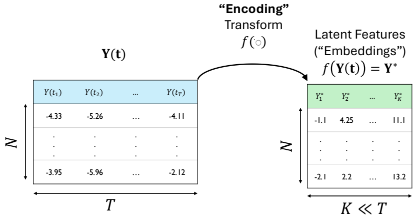
Bayes-FMM-NL

Embed general, potentially **non-linear** encoding $f(\cdot)$ and decoding $g(\cdot)$ transformations into the BayesFMM framework:



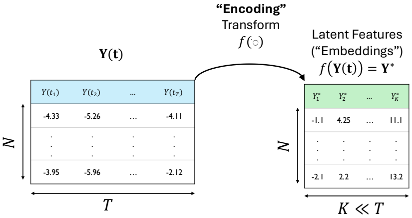
Bayes-FMM-NL

Embed general, potentially **non-linear** encoding $f(\cdot)$ and decoding $g(\cdot)$ transformations into the BayesFMM framework:



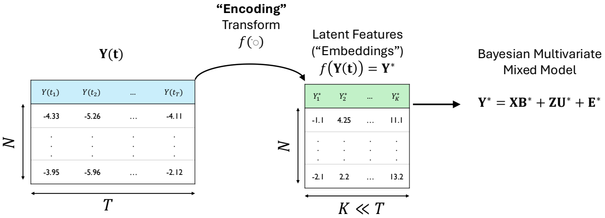
Bayes-FMM-NL

Embed general, potentially **non-linear** encoding $f(\cdot)$ and decoding $g(\cdot)$ transformations into the BayesFMM framework:



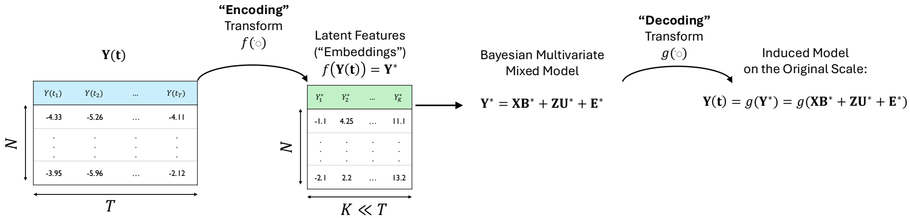
Bayes-FMM-NL

Embed general, potentially **non-linear** encoding $f(\cdot)$ and decoding $g(\cdot)$ transformations into the BayesFMM framework:



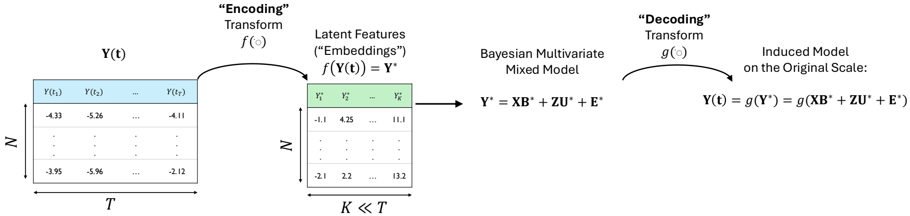
Bayes-FMM-NL

Embed general, potentially **non-linear** encoding $f(\cdot)$ and decoding $g(\cdot)$ transformations into the BayesFMM framework:



Bayes-FMM-NL

Embed general, potentially **non-linear** encoding $f(\cdot)$ and decoding $g(\cdot)$ transformations into the BayesFMM framework:



$\rightarrow$ We will demonstrate our methodology in an application to the MNIST digits dataset.

Example: MNIST Digits Data

MNIST (LeCun, 1998) is a **repeated measures** dataset:

Example: MNIST Digits Data

MNIST (LeCun, 1998) is a **repeated measures** dataset:

1. Writers come from two groups: **high-school students** and **census workers**.

Example: MNIST Digits Data

MNIST (LeCun, 1998) is a **repeated measures** dataset:

1. Writers come from two groups: **high-school students** and **census workers**.
2. Each writer has **multiple replicates of each digit**.

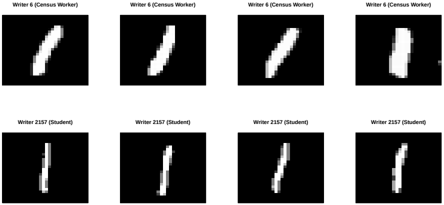
.

Example: MNIST Digits Data

MNIST (LeCun, 1998) is a **repeated measures** dataset:

1. Writers come from two groups: **high-school students** and **census workers**.
2. Each writer has **multiple replicates of each digit**.

.

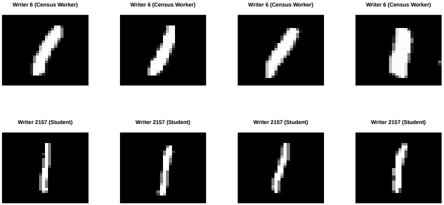


Example: MNIST Digits Data

MNIST (LeCun, 1998) is a **repeated measures** dataset:

1. Writers come from two groups: **high-school students** and **census workers**.
2. Each writer has **multiple replicates of each digit**.

.

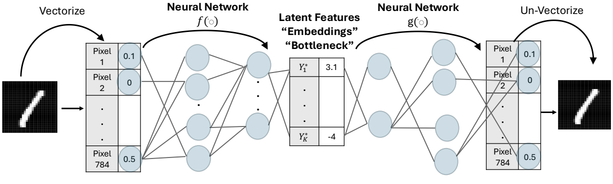


Goal

Regress the 28×28 images on group labels (census vs high-school) while accounting for the multilevel data structure.

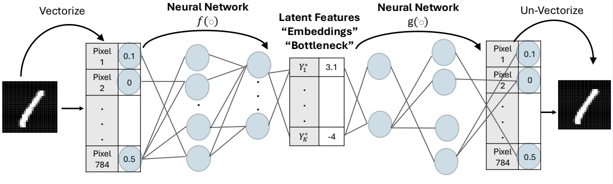
Transformation

Autoencoder

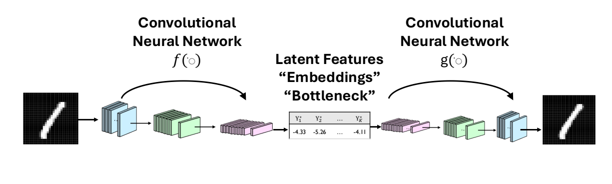


Transformation

Autoencoder



Convolutional Autoencoder



Transformation

Attain reconstruction error < 0.05 for 95^{th} percentile (nearly worst case!):

Figures/compactness-near-losslessness_early_stopping_d

What about the underlying geometry

$\rightarrow$ Latent model assumes $\mathbf{Y}^* = \mathbf{X}\mathbf{B}^* + \mathbf{Z}\mathbf{U}^* + \mathbf{E}^*$, with $\mathbf{U}^*$ and $\mathbf{E}^*$ **multivariate Gaussian**.

What about the underlying geometry

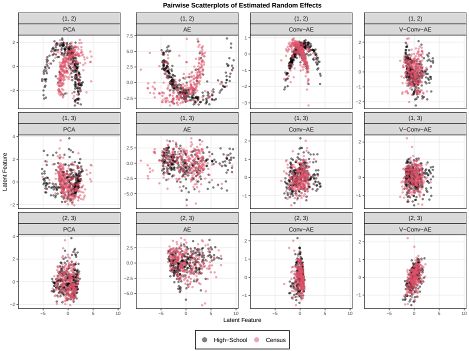
→ Latent model assumes $\mathbf{Y}^* = \mathbf{X}\mathbf{B}^* + \mathbf{Z}\mathbf{U}^* + \mathbf{E}^*$, with $\mathbf{U}^*$ and $\mathbf{E}^*$ **multivariate Gaussian**.

Figures/residual-diagnostics.pdf

What about the underlying geometry

→ Latent model assumes $\mathbf{Y}^* = \mathbf{X}\mathbf{B}^* + \mathbf{Z}\mathbf{U}^* + \mathbf{E}^*$, with $\mathbf{U}^*$ and $\mathbf{E}^*$ **multivariate Gaussian**.

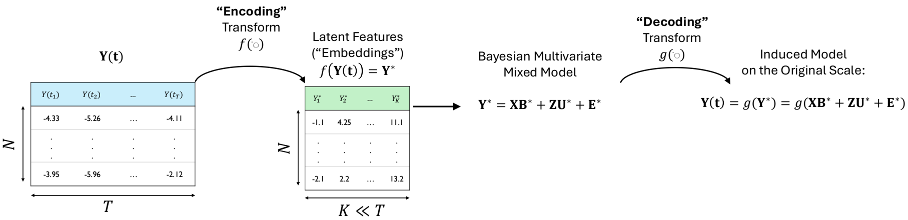
Figures/residual-diagnost



→ Choose **Variational** Convolutional AE (V-Conv-AE) to enforce regularity of latent space!

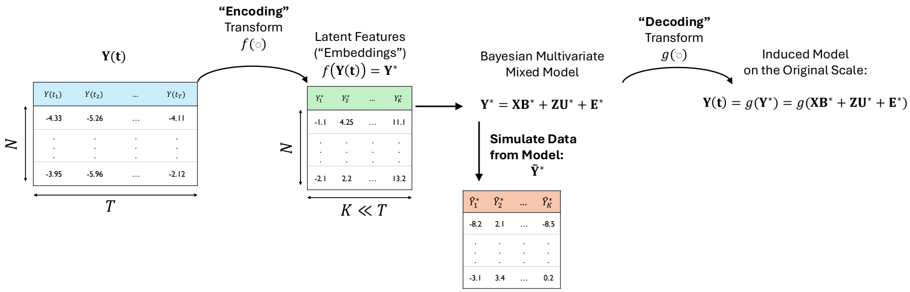
Data-Generative Properties

If encoding and decoding transformations preserve the objects' inherent structure and our latent model is appropriate, then simulations from the model should capture key features of the observed data.



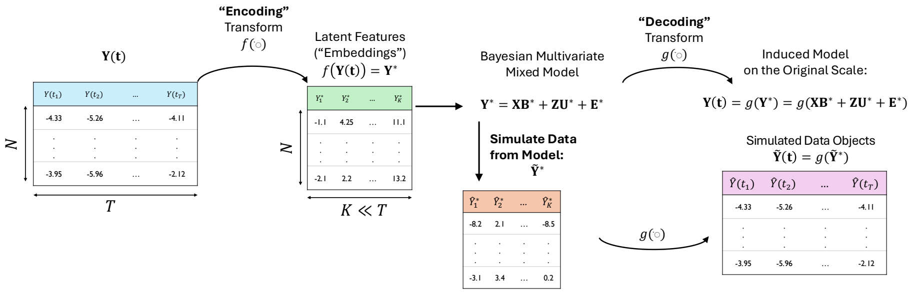
Data-Generative Properties

If encoding and decoding transformations preserve the objects' inherent structure and our latent model is appropriate, then simulations from the model should capture key features of the observed data.



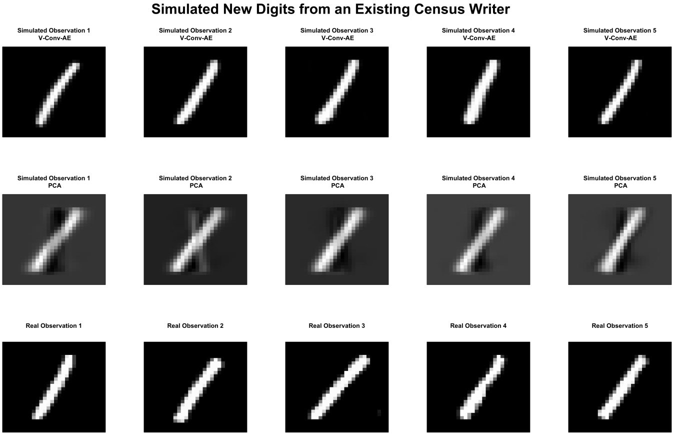
Data-Generative Properties

If encoding and decoding transformations preserve the objects' inherent structure and our latent model is appropriate, then simulations from the model should capture key features of the observed data.

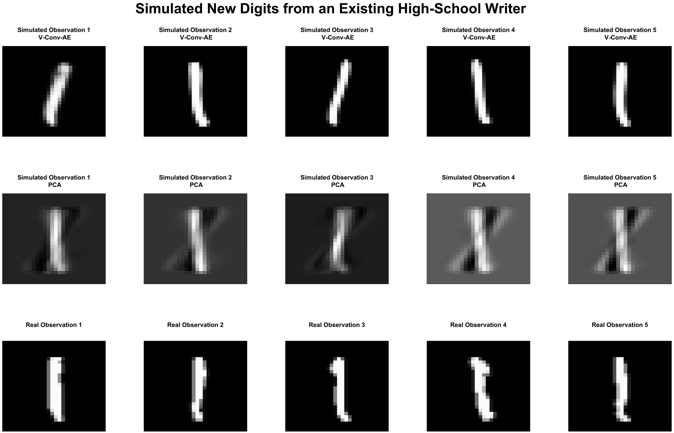


Bayesian approach allows uncertainty in parameter estimates to be propagated through to the generated data.

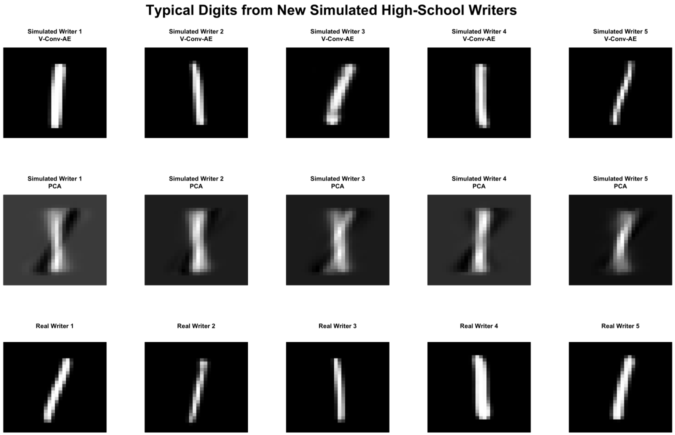
Simulated Data



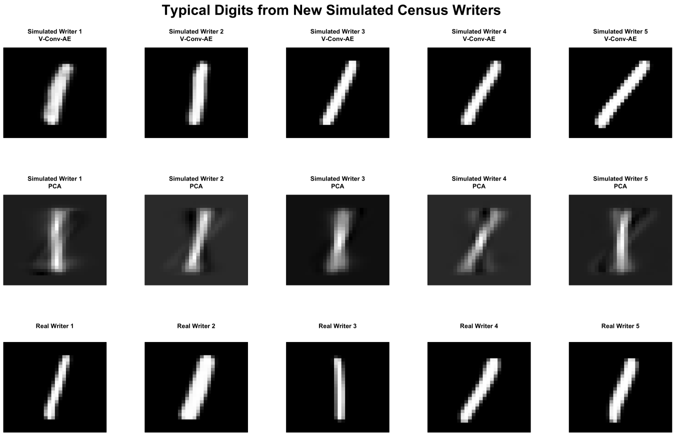
Simulated Data



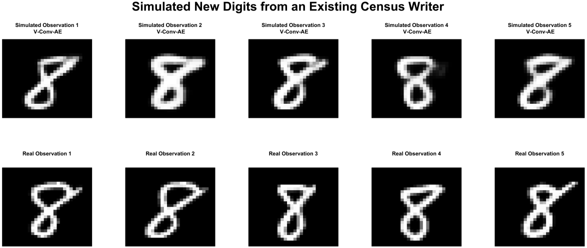
Simulated Data



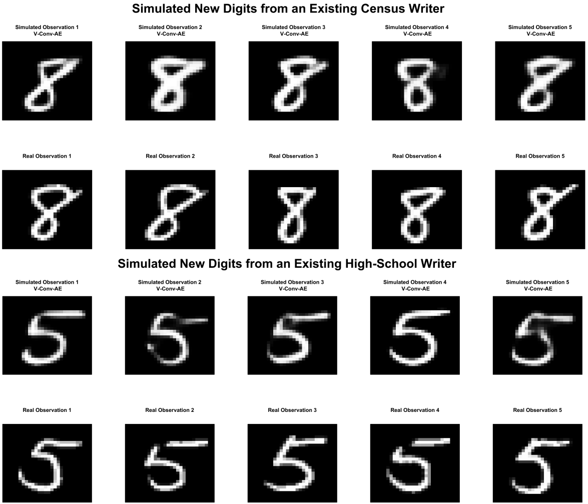
Simulated Data



Simulated Data

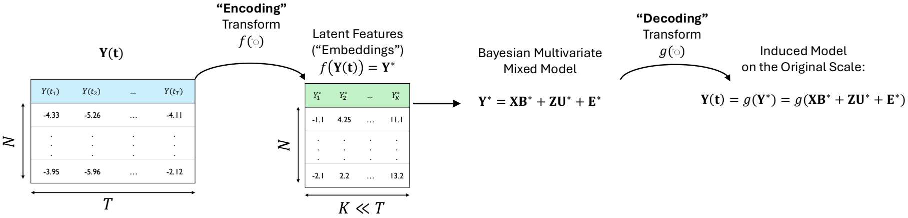


Simulated Data



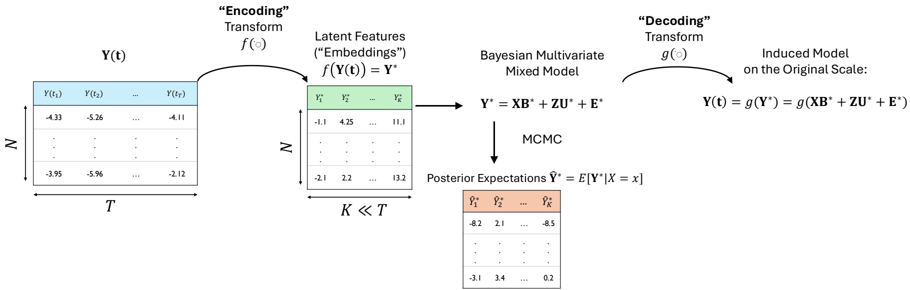
Inference

Since in general $\mathbb{E}\left[g(\mathbf{Y}^*)\right] \neq g\left(\mathbb{E}\left[\mathbf{Y}^*\right]\right)$, we need to think carefully about inference and interpretation:



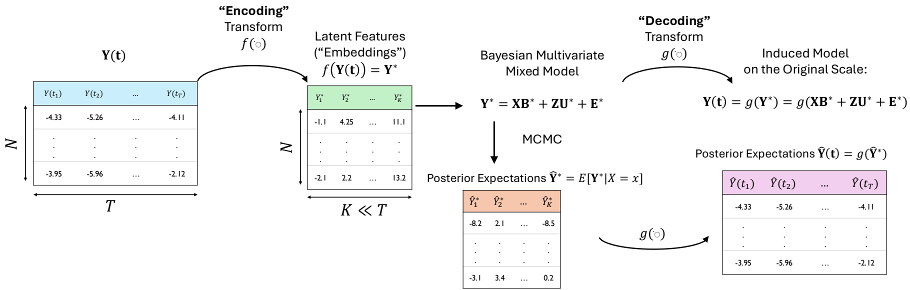
Inference

Since in general $\mathbb{E}[g(\mathbf{Y}^*)] \neq g(\mathbb{E}[\mathbf{Y}^*])$, we need to think carefully about inference and interpretation:



Inference

Since in general $\mathbb{E}\left[g(\mathbf{Y}^*)\right] \neq g\left(\mathbb{E}\left[\mathbf{Y}^*\right]\right)$, we need to think carefully about inference and interpretation:

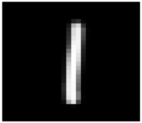


Estimates

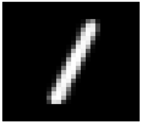
(a) High-School (V-Conv-AE) (b) Census Worker (V-Conv-AE) (c) High-School (PCA) (d) Census Worker (PCA)

Estimates

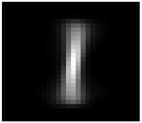
(a) High-School (V-Conv-AE)



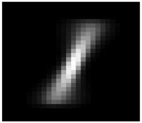
(b) Census Worker (V-Conv-AE)



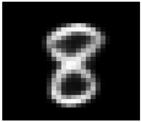
(c) High-School (PCA)



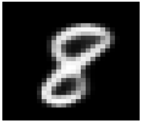
(d) Census Worker (PCA)



(a) High-School (V-Conv-AE)

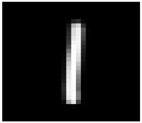


(b) Census Worker (V-Conv-AE)

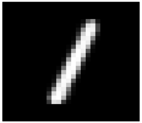


Estimates

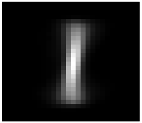
(a) High-School (V-Conv-AE)



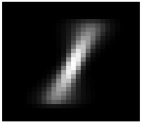
(b) Census Worker (V-Conv-AE)



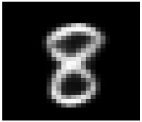
(c) High-School (PCA)



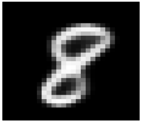
(d) Census Worker (PCA)



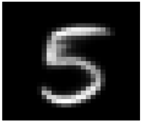
(a) High-School (V-Conv-AE)



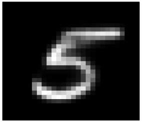
(b) Census Worker (V-Conv-AE)



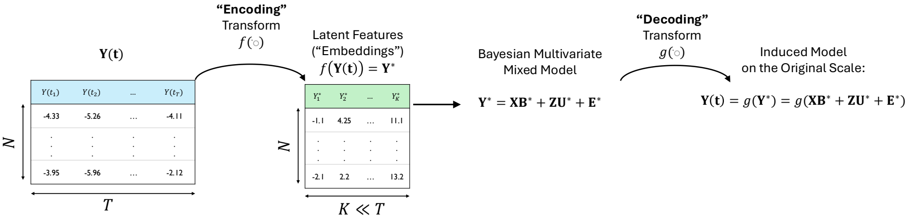
(a) High-School (V-Conv-AE)



(b) Census Worker (V-Conv-AE)

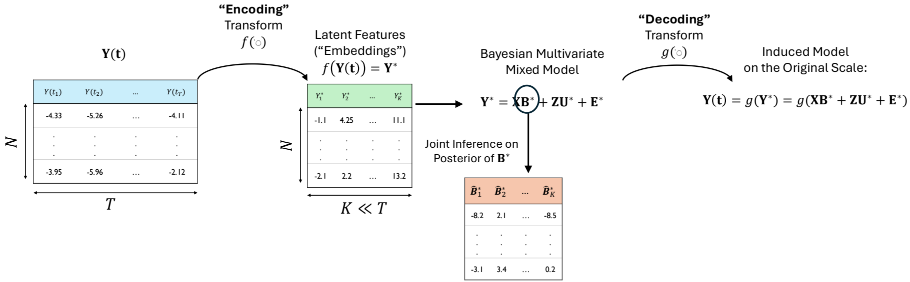


Inference



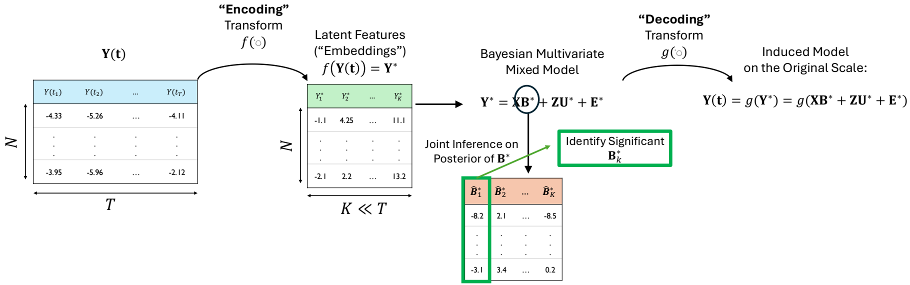
$\mathbf{Y}(\mathbf{t})$
$Y(t_1)$ $Y(t_2)$ ... $Y(t_T)$
-4.33 -5.26 ... -4.11
-3.95 -5.96 ... -2.12
N
T
"Encoding"
Transform
$f(\cdot)$
Latent Features
("Embeddings")
$f\big(\mathbf{Y}(\mathbf{t})\big) = \mathbf{Y}^*$
Y_1^* Y_2^* ... Y_K^*
-1.1 4.25 ... 11.1
-2.1 2.2 ... 13.2
N
$K \ll T$
Bayesian Multivariate
Mixed Model
$\mathbf{Y}^* = \mathbf{X}\mathbf{B}^* + \mathbf{Z}\mathbf{U}^* + \mathbf{E}^*$
"Decoding"
Transform
$g(\cdot)$
Induced Model
on the Original Scale:
$\mathbf{Y}(\mathbf{t}) = g(\mathbf{Y}^*) = g(\mathbf{X}\mathbf{B}^* + \mathbf{Z}\mathbf{U}^* + \mathbf{E}^*)$

Inference



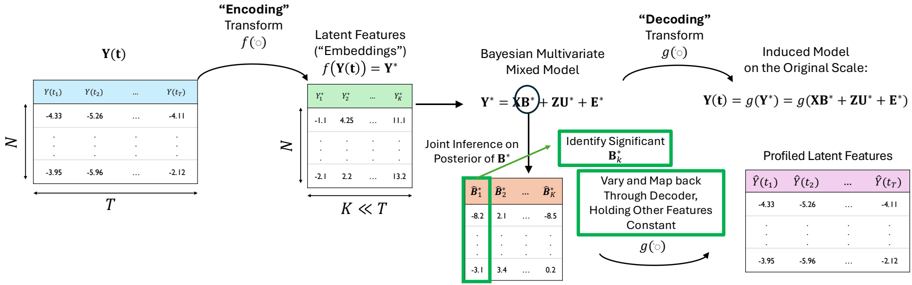
"Encoding"
Transform
$f(\circ)$
$\mathbf{Y}(\mathbf{t})$
$Y(t_1)$ $Y(t_2)$ ... $Y(t_T)$
-4.33 -5.26 ... -4.11
-3.95 -5.96 ... -2.12
N
T
Latent Features
("Embeddings")
$f(\mathbf{Y}(\mathbf{t})) = \mathbf{Y}^*$
Y_1^* Y_2^* ... Y_K^*
-1.1 4.25 ... 11.1
-2.1 2.2 ... 13.2
N
$K \ll T$
Bayesian Multivariate
Mixed Model
$\mathbf{Y}^* = \mathbf{X}\mathbf{B}^* + \mathbf{Z}\mathbf{U}^* + \mathbf{E}^*$
Joint Inference on
Posterior of $\mathbf{B}^*$
$\hat{B}_1^*$ $\hat{B}_2^*$ ... $\hat{B}_K^*$
-8.2 2.1 ... -8.5
-3.1 3.4 ... 0.2
"Decoding"
Transform
$g(\circ)$
Induced Model
on the Original Scale:
$\mathbf{Y}(\mathbf{t}) = g(\mathbf{Y}^*) = g(\mathbf{X}\mathbf{B}^* + \mathbf{Z}\mathbf{U}^* + \mathbf{E}^*)$

Inference

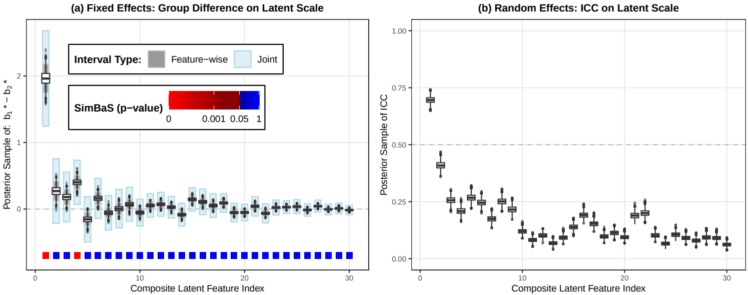


"Encoding"
Transform
$f(\circ)$
$\mathbf{Y}(\mathbf{t})$
Latent Features
("Embeddings")
$f(\mathbf{Y}(\mathbf{t})) = \mathbf{Y}^*$
$Y(t_1)$ $Y(t_2)$ ... $Y(t_T)$
-4.33 -5.26 ... -4.11
-3.95 -5.96 ... -2.12
N
T
Y_1^* Y_2^* ... Y_K^*
-1.1 4.25 ... 11.1
-2.1 2.2 ... 13.2
$K \ll T$
Bayesian Multivariate
Mixed Model
$\mathbf{Y}^* = \mathbf{X}\mathbf{B}^* + \mathbf{Z}\mathbf{U}^* + \mathbf{E}^*$
"Decoding"
Transform
$g(\circ)$
Induced Model
on the Original Scale:
$\mathbf{Y}(\mathbf{t}) = g(\mathbf{Y}^*) = g(\mathbf{X}\mathbf{B}^* + \mathbf{Z}\mathbf{U}^* + \mathbf{E}^*)$
Joint Inference on
Posterior of $\mathbf{B}^*$
Identify Significant
$\mathbf{B}_k^*$
$\widehat{B}_1^*$ $\widehat{B}_2^*$ ... $\widehat{B}_K^*$
-8.2 2.1 ... -8.5
-3.1 3.4 ... 0.2

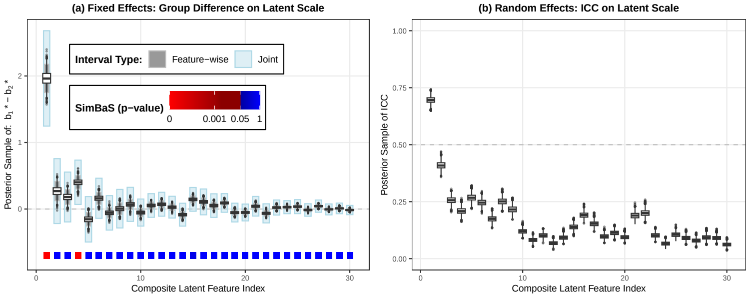
Inference


“Encoding”
Transform
$f(\cdot)$
$\mathbf{Y}(\mathbf{t})$
$Y(t_1)$ $Y(t_2)$ ... $Y(t_T)$
-4.33 -5.26 ... -4.11
-3.95 -5.96 ... -2.12
N
T
Latent Features
(“Embeddings”)
$f\big(\mathbf{Y}(\mathbf{t})\big) = \mathbf{Y}^*$
Y_1^* Y_2^* ... Y_K^*
-1.1 4.25 ... 11.1
-2.1 2.2 ... 13.2
$K \ll T$
Bayesian Multivariate
Mixed Model
$\mathbf{Y}^* = \mathbf{X}\mathbf{B}^* + \mathbf{Z}\mathbf{U}^* + \mathbf{E}^*$
Joint Inference on
Posterior of $\mathbf{B}^*$
$\hat{B}_1^*$ $\hat{B}_2^*$ ... $\hat{B}_K^*$
-8.2 2.1 ... -8.5
-3.1 3.4 ... 0.2
Identify Significant
$\mathbf{B}_k^*$
Vary and Map back
Through Decoder,
Holding Other Features
Constant
$g(\cdot)$
“Decoding”
Transform
$g(\cdot)$
Induced Model
on the Original Scale:
$\mathbf{Y}(\mathbf{t}) = g(\mathbf{Y}^*) = g(\mathbf{X}\mathbf{B}^* + \mathbf{Z}\mathbf{U}^* + \mathbf{E}^*)$
Profiled Latent Features
$\hat{Y}(t_1)$ $\hat{Y}(t_2)$ ... $\hat{Y}(t_T)$
-4.33 -5.26 ... -4.11
-3.95 -5.96 ... -2.12

Inference

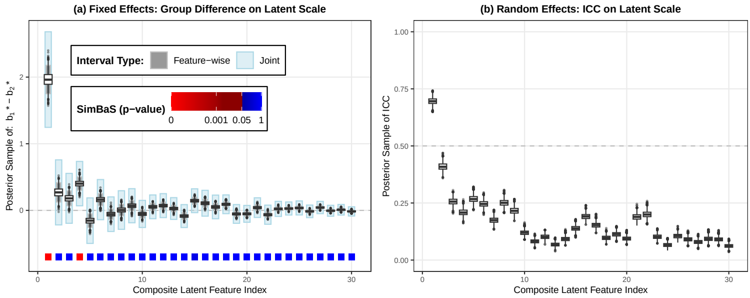

(a) Fixed Effects: Group Difference on Latent Scale
Interval Type: Feature−wise Joint
SimBaS (p−value)
0 0.001 0.05 1
Posterior Sample of: $b_1* - b_2*$
Composite Latent Feature Index
(b) Random Effects: ICC on Latent Scale
Posterior Sample of ICC
Composite Latent Feature Index

Inference



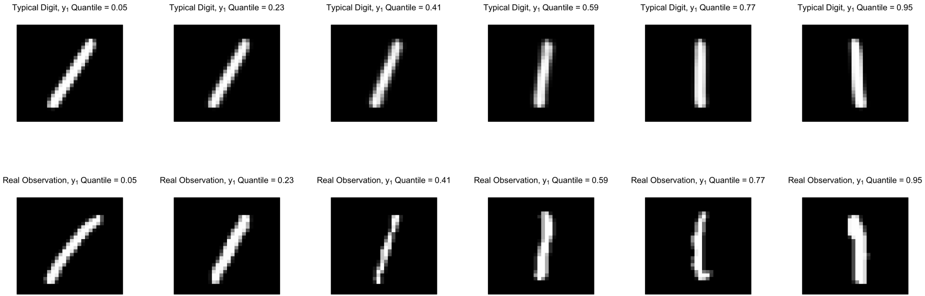
→ Identify Feature 1 as a driver of **between-group** differences.

Inference



→ Identify Feature 1 as a driver of **between-group** differences.
→ Higher-order features important for **between-subject** variation.

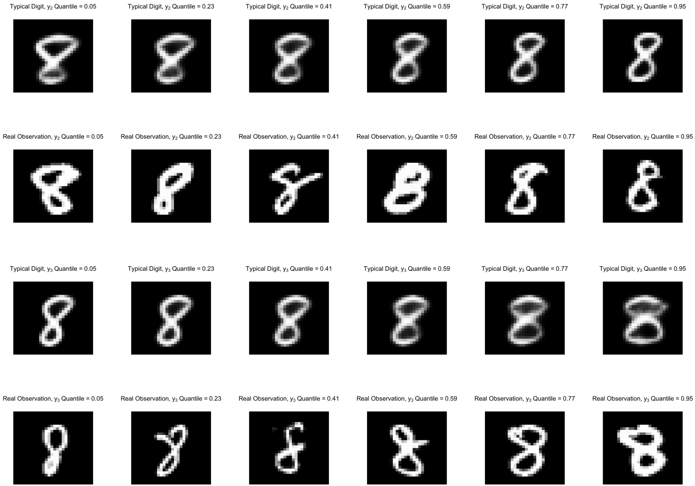
Inference


Typical Digit, y1 Quantile = 0.05
Typical Digit, y1 Quantile = 0.23
Typical Digit, y1 Quantile = 0.41
Typical Digit, y1 Quantile = 0.59
Typical Digit, y1 Quantile = 0.77
Typical Digit, y1 Quantile = 0.95
Real Observation, y1 Quantile = 0.05
Real Observation, y1 Quantile = 0.23
Real Observation, y1 Quantile = 0.41
Real Observation, y1 Quantile = 0.59
Real Observation, y1 Quantile = 0.77
Real Observation, y1 Quantile = 0.95

Inference

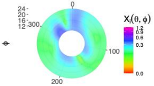
Figures/combined-profile-features-2-3-digit_1.png

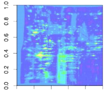
Inference

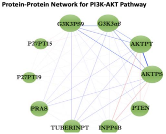

Typical Digit, y_2 Quantile = 0.05
Typical Digit, y_2 Quantile = 0.23
Typical Digit, y_2 Quantile = 0.41
Typical Digit, y_2 Quantile = 0.59
Typical Digit, y_2 Quantile = 0.77
Typical Digit, y_2 Quantile = 0.95
Real Observation, y_2 Quantile = 0.05
Real Observation, y_2 Quantile = 0.23
Real Observation, y_2 Quantile = 0.41
Real Observation, y_2 Quantile = 0.59
Real Observation, y_2 Quantile = 0.77
Real Observation, y_2 Quantile = 0.95
Typical Digit, y_3 Quantile = 0.05
Typical Digit, y_3 Quantile = 0.23
Typical Digit, y_3 Quantile = 0.41
Typical Digit, y_3 Quantile = 0.59
Typical Digit, y_3 Quantile = 0.77
Typical Digit, y_3 Quantile = 0.95
Real Observation, y_3 Quantile = 0.05
Real Observation, y_3 Quantile = 0.23
Real Observation, y_3 Quantile = 0.41
Real Observation, y_3 Quantile = 0.59
Real Observation, y_3 Quantile = 0.77
Real Observation, y_3 Quantile = 0.95

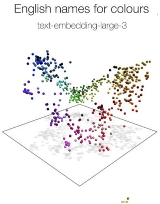
A General Framework

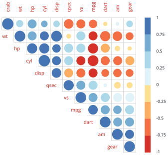
Data Objects





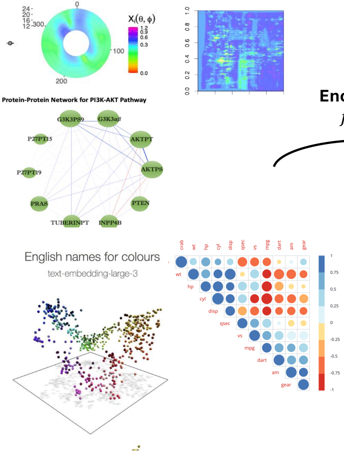






A General Framework

Data Objects



Encoding
$f(\cdot)$

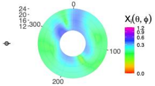
Lower-Dimensional Space of Latent Features

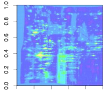
Y_1^*	Y_2^*	...	Y_K^*
-4.33	-5.26	...	-4.11

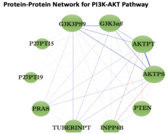
Amenable to Familiar Statistical Modeling
E.g., Regression, Clustering, Classification

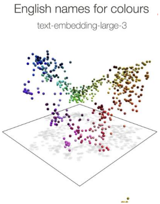
A General Framework

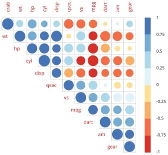
Data Objects











Encoding
$f(\cdot)$

Lower-Dimensional Space of Latent Features

Y_1^*	Y_2^*	...	Y_K^*
-4.33	-5.26	...	-4.11

Amenable to Familiar Statistical Modeling
E.g., Regression, Clustering, Classification

Decoding
$g(\cdot)$

Original Data-Object Scale

- Inference
- Interpretation
- Data Generation

A General Framework

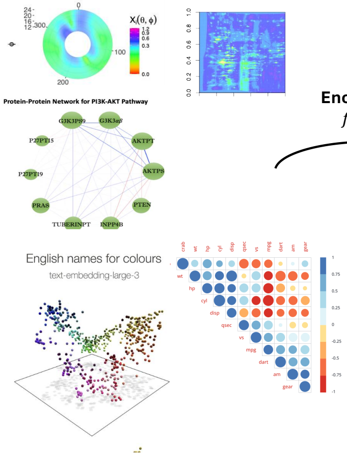
- Transformation functions $f(\cdot)$ and $g(\cdot)$ can:
  - Utilize modern machine-learning tools.
  - Incorporate analytic transformations that respect the objects' geometry.
  - Induce favorable latent space geometry.
  - Use compositions, e.g., $f(\cdot) = f_2 \circ f_1(\cdot)$
- Framework allows us to:

A General Framework

- Transformation functions $f(\cdot)$ and $g(\cdot)$ can:
  - Utilize modern machine-learning tools.
  - Incorporate analytic transformations that respect the objects' geometry.
  - Induce favorable latent space geometry.
  - Use compositions, e.g., $f(\cdot) = f_2 \circ f_1(\cdot)$
- Framework allows us to:
  - Model modern data objects using familiar statistical tools $\rightarrow$ principled approaches to estimation, inference, hypothesis testing etc.,
  - Leverage the power of contemporary approaches for data representation within familiar and well-tested statistical modeling frameworks.

Thank You For Listening

Data Objects



Encoding
$f(\cdot)$

Lower-Dimensional Space of Latent Features

Y_1^*	Y_2^*	...	Y_K^*
-4.33	-5.26	...	-4.11

Amenable to Familiar Statistical Modeling
E.g., Regression, Clustering, Classification

Decoding
$g(\cdot)$

Original Data-Object Scale

- Inference
- Interpretation
- Data Generation