Bayesian Estimation of Physical Activity from Wearable Sensor Measurements

Jiwon Lee

University of Cincinnati

August 26, 2024

Content

Overview

- Accelerometer sensor-based wearable device.
- Acceleration (body movement) is measured in standard gravity units (G, 9.8 m/s^2), typically in three axes (X, Y, Z).
- Used to infer physical activities like walking, running, sitting, or sleeping.

Actigraph wGT3X-BT: tri-axial accelerometry-based activity monitor. source: https://theactigraph.com/actigraph-wgt3x-bt

Data

- The raw data from an accelerometer is typically a time series of acceleration values along each axis.

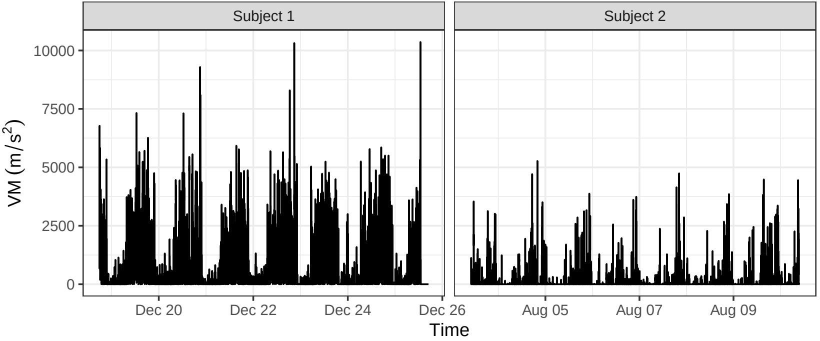
A snapshot of the sample dataset from a study participant

Time Stamp	axis1	axis2	axis3	VM
2017-09-06 15:30:00	116	0	56	129
2017-09-06 15:31:00	126	32	218	254
2017-09-06 15:32:00	845	663	1213	1621

Vector Magnitude (VM) $= \sqrt{\text{axis1}^2 + \text{axis2}^2 + \text{axis3}^2}$.

- Use VM as a primary outcome.

Raw Signal Data



Physical activity patterns vary by subjects.

Common Practice

Physical activity analysis is conducted using the following procedure:

1. Sensors collect raw signal data.
2. Preprocessing (depend on **device-specific pre-defined algorithms**)
    - Handling missing data.
    - Filtering (smoothing) noise.
    - Dividing continuous data into meaningful segments.
        - choosing a time window (30 seconds or 1 minutes, etc.)
        - **separating periods of activity from rest**

Common Practice

3. Feature Extraction
   - Common features: simple summary statistics
     e.g, mean acceleration and variance.
   - ActiGraph wGT3X-BT provides:
     - Total Movement
     - Moderate to Vigorous Physical Activity (MVPA)
     - Non-Sedentary Time
     - Step Count
     - Energy Expenditure
4. Activity Recognition
   - Use machine learning or deep learning models to classify the physical activities based on the extracted features.

Common Practice

5. Monitoring and Feedback
   Based on the previous steps, practitioners can figure out
   - Whether there is any change in daily-routine or sleep pattern.
   - Whether a subject reaches a daily step goal.
   - Whether a subject needs to take frequent breaks:
     e.g. sedentary for long periods.

What needs improvement

- Heuristically designed algorithms that process the outcomes.
  - E.g. step counts = number of times a measurement exceeds a **threshold**.
- Lack of literature to process and analyze raw datasets.
  - Most of the existing methodology are primarily for validation studies.
- Outcomes associated with additional information from wearers, or pre-obtained data in the laboratory.
  - Participants visit a facility in advance to obtain a training data for activity classification.

Additional Information - Example

TWO WEEK SLEEP DIARY

AASM | SLEEP EDUCATION

INSTRUCTIONS:
(1) Write the date, day of the week, and type of day: Work, School, Day Off, or Vacation. **(2)** Put the letter “C” in the box when you have coffee, cola or tea. Put “M” when you take any medicine. Put “A” when you drink alcohol. Put “E” when you exercise. **(3)** Put a “B” in the box to show when you go to bed. Put a “Z” in the box that shows when you think you fell asleep. **(4)** Put a “Z” in all the boxes that show when you are asleep at night or when you take a nap during the day. **(5)** Leave boxes empty to show when you wake up at night and when you are awake during the day.

SAMPLE ENTRY BELOW: On a Monday when I worked, I jogged on my lunch break at 1 PM, had a glass of wine with dinner at 6 PM, fell asleep watching TV from 7 to 8 PM, went to bed at 10:30 PM, fell asleep around Midnight, woke up and couldn't got back to sleep at about 4 AM, went back to sleep from 5 to 7 AM, and had coffee and medicine at 7 AM.

Date	Day of the week	Type of Day (Work, School, Day Off, Vacation)	Noon	1 PM	2 PM	3 PM	4 PM	5 PM	6 PM	7 PM	8 PM	9 PM	10 PM	11 PM	Midnight	1 AM	2 AM	3 AM	4 AM	5 AM	6 AM	7 AM	8 AM	9 AM	10 AM	11 AM
sample	Mon.	Work		E					A	Z			B		Z	Z	Z	Z		Z	Z	C M				

week 1

Sleep diary to get sleep-wake information.

Research Goals

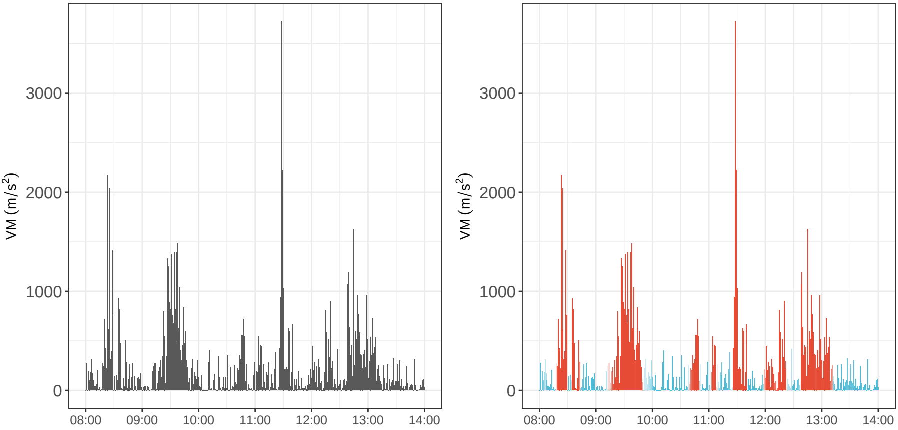
- **Data-driven** algorithm to process raw signal data.
  - Separating periods of activity from rest
  - A manufacturer-provided single threshold is not applicable to all populations.
- **Unsupervised** algorithm without additional information.
  - Leverage the advantage of non-intrusive and continuous data collection.

Content

State Labeling

- Assume that there are two states alternating: **active** and **resting**.
- Identify the active states to derive meaningful metrics for physical activity.
- In the active state, **dynamic physical movements** are expected, such as walking and running.
- The resting state comprises sleep and resting periods with **little body movement**.
- These two states are not observed; they are the **latent (hidden)** states.

State Labeling

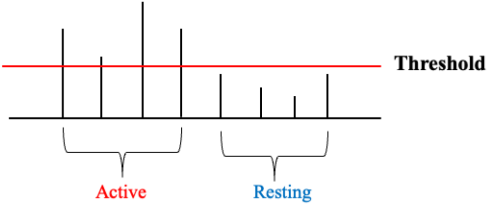


The figure on the right is what we obtained using our proposed method.

Existing Methods and Limitations

1. Heuristic methods

   A measurement VM is labeled as "active" if VM $>$ threshold.



2. The **hidden Markov model** is a popular model in the hidden state inference problem.
   - Its strong Markovian assumption is not suitable for wearable devices measurements which often have long-range dependence.

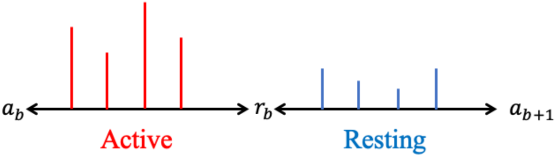
Existing Methods and Limitations

3. The **conditional random fields model**, an extension of hidden Markov models with less strict Markovian assumption using the features extracted from data (Lafferty et al., 2001).
    - Limited features to use, e.g., mean or variance of measurements given a window.

$\rightarrow$ These models should be trained in a **supervised manner**, and hence require additional sources of information.

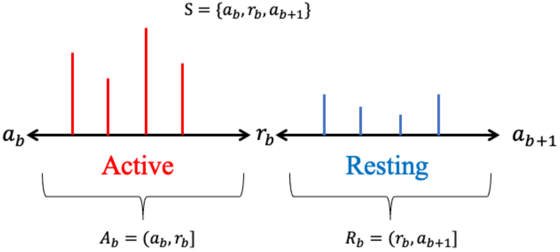
Proposed Methods

- An **interval-based state labeling** inspired by the **change-point detection** methodology of Green (1995).
- A pair of points, the start and end points, constitutes an interval labeled as either an active or a resting state.
- Each measurement between these points is then assigned the same state label, depending on the interval to which it belongs.



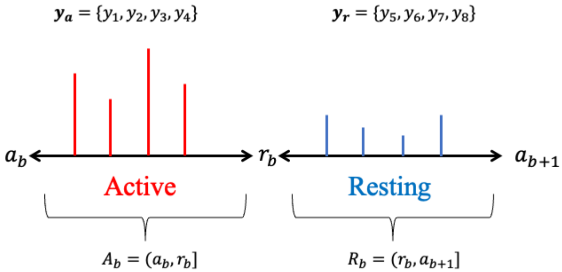
Multiple change points mark the transitions between the two states.

Notations



- $\mathcal{S} = \{a_1, r_1, \ldots, a_B, r_B\}$: the set of change points with the specific ordering $r_{b-1} < a_b < r_b < a_{b+1}$.
- $A_b = (a_b, r_b]$ and $R_b = (r_b, a_{b+1}]$: b-th active and resting intervals, respectively.

Notations



- y_t: the observed value at time $t \in \mathcal{T}^y$ where $\mathcal{T}^y = \{1, \ldots, T\}$.
- $\boldsymbol{y}_a = \{y_t : t \in A_b,\ b = 1, \ldots, B\}$: the collection of measurements in active states.
  $\boldsymbol{y}_r = \{y_t : t \in R_b,\ b = 1, \ldots, B-1\}$: the collection of measurements in resting states.

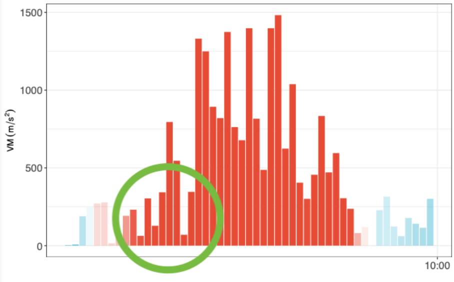
Model for Active State Measurements

$$f_A(\boldsymbol{y}_a) \sim N(\mu\mathbf{1}, \sigma^2\boldsymbol{\Sigma}),$$

$$\boldsymbol{\Sigma} = \begin{cases} \exp\left(-\frac{|t-t'|}{l}\right) + g\delta_{t,t'}, & \text{for diagonal elements} \\ \exp\left(-\frac{|t-t'|}{l}\right), & \text{for off-diagonal elements} \end{cases}$$

with the nugget effect g fixed as a small number called jitter (Neal, 1999).

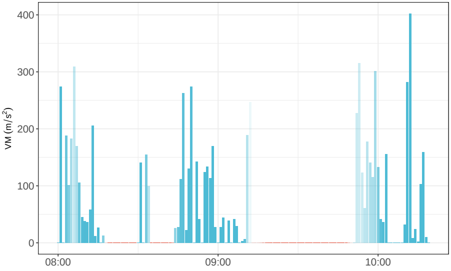
Model for Active State Measurements



Use of the temporal kernel component

- A few abrupt drops in the measurements **do not immediately trigger a change** in the state from active to resting.
- The transition process less sensitive than the manufacturer-setting threshold for the measurement level, ensuring **the state labeling is robust**.

Model for Resting State Measurements



- Consider the following characteristics of the measurements recorded during the resting states $\boldsymbol{y}_r$:

  (a) having excessive zeros

  (b) being heavily skewed to the right

  (c) having sporadic spikes

Model for Resting State Measurements

- A Poisson distribution

  $\rightarrow$ not flexible enough to handle spikes occurring at varying scales.
- A mixture of Poisson distributions for $\boldsymbol{y}_r$ with K mixture components:

  $\rightarrow$ For $y_t \in \boldsymbol{y}_r$: $f_R(y_t) = \sum_{k=1}^{K} \pi_k \text{Pois}(y_t; \lambda_k)$,
  where $\pi_k = p(z_t = k)$ for $k = 1, \ldots, K$ and z_t represents a membership indicator of the mixture components.

Justification of the underlying distributions

- Poisson distribution:

  – A measurement on each axis represents the number of acceleration that exceed a threshold.

  – Use $\text{VM} = \sqrt{\text{axis1}^2 + \text{axis2}^2 + \text{axis3}^2}$ then rounded, but it still retains the characteristics of count data.

- Normal distribution:

  – Rounding to the integers has minimal impact on the distribution due to the large sample size.

Prior Distributions (for model parameters)

Parameters of interest: $\Theta = \left\{\mu, \sigma^2, l\right\}$, $\Phi = \{\boldsymbol{\lambda}, \boldsymbol{\pi}, \mathcal{Z}\}$, and $\mathcal{S} = \{a_1, \ldots, r_B\}$.

1. For $\boldsymbol{y}_a$; $\Theta = \left\{\mu, \sigma^2, l\right\}$

$$\boldsymbol{y}_a \sim N\left(\mu\mathbf{1}, \sigma^2 \exp\left(-\frac{|t-t'|}{l}\right) + g\delta_{t,t'}\right).$$

   – μ: flat prior

   – l and σ^2: uniform priors

2. For $\boldsymbol{y}_r$; $\Phi = \{\boldsymbol{\lambda}, \boldsymbol{\pi}, \mathcal{Z}\}$

$$f_R(y_t) = \sum_{k=1}^{K} \pi_k \text{Pois}(y_t; \lambda_k)$$

   – A Dirichlet process prior for a mixing measure G with stick-breaking representation.

Prior Distributions (for model parameters)

Parameters of interest: $\Theta = \left\{\mu, \sigma^2, l\right\}$, $\Phi = \{\boldsymbol{\lambda}, \boldsymbol{\pi}, \mathcal{Z}\}$, and $\mathcal{S} = \{a_1, \ldots, r_B\}$.

2. For $\boldsymbol{y}_r$; $\Phi = \{\boldsymbol{\lambda}, \boldsymbol{\pi}, \mathcal{Z}\}$

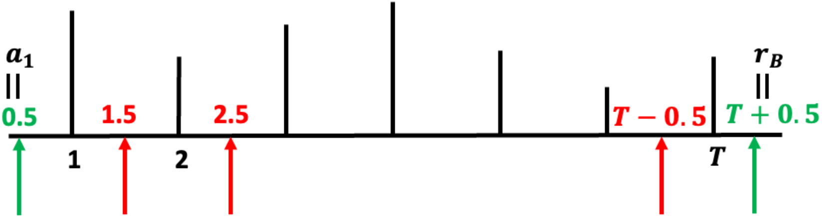
$$f_R(y_t) = \sum_{k=1}^{K} \pi_k \text{Pois}(y_t; \lambda_k)$$

– $G = \sum_{k=1}^{K} \pi_k \delta_{\lambda_k}$, $\lambda_k \sim G_0$ i.i.d (Ishwaran and James, 2001).

– $\pi_k = U_k \prod_{j=1}^{k-1}(1 - U_j)$ for $k = 2, \ldots, K-1$ and $\pi_1 = U_1$ where $U_j \sim \text{Beta}(1, \alpha)$ for $j = 1, \ldots, K-1$ and $U_K = 1$.

Prior Distributions (for change points)

3. For $\mathcal{S} = \{a_1, \ldots, r_B\}$ (fix a_1 and r_B)
   - $2B - 2$ order statistics of $(r_1, \ldots, a_B) \sim U[1.5, T - 0.5]$.
   - $\pi(\mathcal{S}) = \frac{(2B-2)!}{(T-1)^{(2B-2)}}$.



where $\mathcal{S} \in \{1.5, \ldots, T - 0.5\}$.

Reversible Jump MCMC

- Inference about the set of change points $\mathcal{S}$:

  – True number of the change points, i.e., **the dimension of $\mathcal{S}$ is unknown**.

  – Ordinary MCMC cannot handle the varying dimension of parameter spaces.

- Reversible Jump MCMC (RJMCMC) (Green 1995)

  – allows a Markov chain to traverse parameter spaces of different dimensions.

RJMCMC - General Framework

- Ordinary MCMC: acceptance probability for the proposed parameter θ' given the model $\mathcal{M}$:

$$\min\left\{1,\ \frac{p(\theta'|\boldsymbol{y},\mathcal{M})q(\theta|\theta')}{p(\theta|\boldsymbol{y},\mathcal{M})q(\theta'|\theta)}\right\}.$$

- RJMCMC: acceptance probability includes Jacobian

$$\min\left\{1,\ \frac{p(\theta'|\boldsymbol{y},\mathcal{M}')\boldsymbol{q}\left[(\boldsymbol{\mathcal{M}},\boldsymbol{\theta})|(\boldsymbol{\mathcal{M}'},\boldsymbol{\theta'})\right]}{p(\theta|\boldsymbol{y},\mathcal{M})\boldsymbol{q}\left[(\boldsymbol{\mathcal{M}'},\boldsymbol{\theta'})|(\boldsymbol{\mathcal{M}},\boldsymbol{\theta})\right]}\left|\frac{\boldsymbol{\partial(\theta')}}{\boldsymbol{\partial(\theta,u)}}\right|\right\}.$$

- Probability for proposing $\mathcal{M}'$ and Jacobian term are added.

RJMCMC - General Framework

- The corresponding parameters θ' and θ to $\mathcal{M}'$ and $\mathcal{M}$, respectively:

$$\text{dim}(\theta') \neq \text{dim}(\theta).$$

- Auxiliary parameter $u \sim g(\cdot)$ to match the dimension

$$\text{dim}(\theta') = \text{dim}(\theta, u).$$

- Set $\theta' = h(\theta, u)$ for a deterministic function $h(\cdot)$.
  - E.g. $\text{dim}(\theta') = 2$ and $\text{dim}(\theta) = 1$, then draw $u \sim g(\cdot)$ and set:

$$\begin{aligned} \theta' &= (\theta'_1, \theta'_2) \\ \theta'_1 &= \theta + u \\ \theta'_2 &= \theta - u. \end{aligned}$$

Reverse move is then $\theta = (\theta'_1 + \theta'_2)/2$.

RJMCMC - General Framework

- The acceptance probability for RJMCMC:

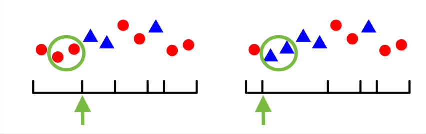
$$\min\left\{1,\ \frac{p(\theta'|\boldsymbol{y},\mathcal{M}')q\left[(\mathcal{M},\theta)|(\mathcal{M}',\theta')\right]}{p(\theta|\boldsymbol{y},\mathcal{M})q\left[(\mathcal{M}',\theta')|(\mathcal{M},\theta)\right]}\left|\frac{\partial(\theta')}{\partial(\theta,u)}\right|\right\}$$

where $q\left[(\mathcal{M}',\theta')|(\mathcal{M},\theta)\right] = j(\mathcal{M}'|\mathcal{M})g(u)$.

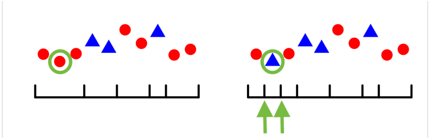
– Jacobian depends on the function $h(\cdot)$.

RJMCMC - Move Types

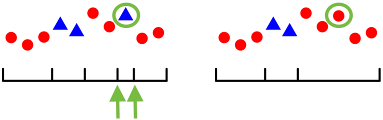
1. Update - change the locations of $r_1, \ldots, a_B \in \mathcal{S}$ (p_{update}).



2. Birth - create a new interval (p_{birth}).

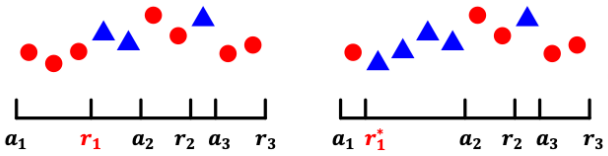


3. Death - remove an existing interval (p_{death}).



RJMCMC - Update Move

Change the location of $2B - 2$ elements in $\mathcal{S} = \{r_1, \ldots, a_B\}$ under the restriction: $a_b < r_b < a_{b+1}$.

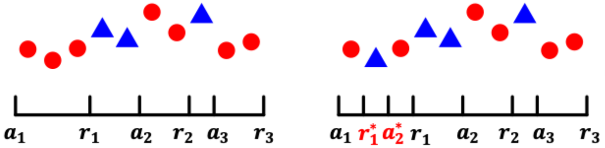


Update r_b (or a_b):

1) Propose $r_b' \sim U[a_b + 1, a_{b+1} - 1]$
2) Accept r_b' with $\alpha_u = \min\{1, \gamma_u\}$.
$\gamma_u = [f_A(\boldsymbol{y}_a')f_R(\boldsymbol{y}_r')] \,/\, [f_A(\boldsymbol{y}_a)f_R(\boldsymbol{y}_r)]$ where
$\boldsymbol{y}_a' = \{y_t : a_b < t \leq r_b'\}$, $\boldsymbol{y}_r' = \{y_t : r_b' < t \leq a_{b+1}\}$,
$\boldsymbol{y}_a = \{y_t : a_b < t \leq r_b\}$, and $\boldsymbol{y}_r = \{y_t : r_b < t \leq a_{b+1}\}$.

RJMCMC - Birth Move

Create a new interval $(a'_q, r'_q]$ or $(r'_q, a'_{q+1}]$.



Birth of a resting interval $(r'_q, a'_{q+1}]$.

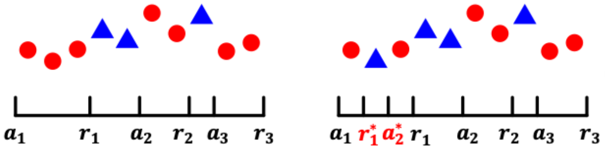
1) Choose $(a_q, r_q]$ where $(r'_q, a'_{q+1}]$ is proposed.

$$\mathcal{S} = \{a_1, \ldots, a_q, \quad , \quad , r_q, \ldots, r_B\}$$
$$\mathcal{S}' = \{a_1, \ldots, a_q, r'_q, a'_{q+1}, r_q, \ldots, r_B\}$$

$$\dim(\mathcal{S}') = \dim(\mathcal{S}) + 2.$$

RJMCMC - Birth Move

Create a new interval $(a'_q, r'_q]$ or $(r'_q, a'_{q+1}]$.



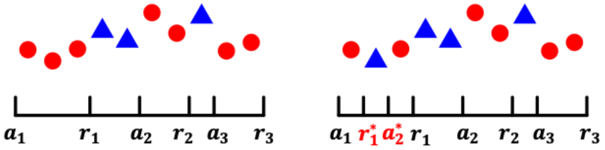
Involved probability:

$(a_q, r_q]$ is randomly chosen **proportional** to its length:

$$\pi_q^A := \frac{(r_q - a_q)}{\sum_{b=1}^{B}(r_b - a_b)}.$$

RJMCMC - Birth Move

Create a new interval $(a'_q, r'_q]$ or $(r'_q, a'_{q+1}]$.



$\mathcal{S}' = \{a_1, \ldots, a_q, r'_q, a'_{q+1}, r_q, \ldots, r_B\}$

2) Propose u_1 and u_2 to match the dimension.
   u_1, u_2: order statistics from $U[a_q + 1, r_q - 1]$,

$$g(u_1, u_2) = 2! \frac{1}{(r_q - a_q - 1)^2}.$$

3) Use identity function to link (u_1, u_2) to (r'_q, a'_{q+1}).

$$u_1 = r'_q \quad \text{and} \quad u_2 = a'_{q+1}.$$

RJMCMC - Birth Move

Break down the acceptance probability for $(r'_q, a'_{q+1}]$:

$$\min\left\{1,\ \frac{p(\theta'|\boldsymbol{y},\mathcal{M}')q\left[(\mathcal{M},\theta)|(\mathcal{M}',\theta')\right]}{p(\theta|\boldsymbol{y},\mathcal{M})q\left[(\mathcal{M}',\theta')|(\mathcal{M},\theta)\right]}\left|\frac{\partial(\theta')}{\partial(\theta,u)}\right|\right\}.$$

Posterior ratio $p(\theta'|\boldsymbol{y},\mathcal{M}')/p(\theta|\boldsymbol{y},\mathcal{M})$:

- $p(\theta'|\boldsymbol{y},\mathcal{M}')$

$$\begin{aligned} p(\theta'|\boldsymbol{y},\mathcal{M}') &= p(\boldsymbol{y}|\theta',\mathcal{M}')\pi(\theta') \\ &= f_R(\boldsymbol{y}_q)\pi(\mathcal{S}') \\ &= f_R(\boldsymbol{y}_q)\frac{(2B)!}{(T-1)^{2B}} \end{aligned}$$

  where $\boldsymbol{y}_q = \{y_t : r'_q < y_t < a'_{q+1}\}$.

- $p(\theta|\boldsymbol{y},\mathcal{M}) = f_A(\boldsymbol{y}_q)\pi(\mathcal{S}) = f_A(\boldsymbol{y}_q)\frac{(2B-2)!}{(T-1)^{2B-2}}$.

RJMCMC - Birth Move

Break down the acceptance probability for $(r'_q, a'_{q+1}]$:

$$\min\left\{1,\ \frac{p(\theta'|\boldsymbol{y},\mathcal{M}')q\left[(\mathcal{M},\theta)|(\mathcal{M}',\theta')\right]}{p(\theta|\boldsymbol{y},\mathcal{M})q\left[(\mathcal{M}',\theta')|(\mathcal{M},\theta)\right]}\left|\frac{\partial(\theta')}{\partial(\theta,u)}\right|\right\}.$$

Proposal ratio $q\left[(\mathcal{M},\theta)|(\mathcal{M}',\theta')\right]/q\left[(\mathcal{M}',\theta')|(\mathcal{M},\theta)\right]$:

- $q\left[(\mathcal{M},\theta)|(\mathcal{M}',\theta')\right]$ requires a reverse move (death)

$$\begin{aligned}\mathcal{S}' &= \{a_1,\ldots,a_q,{\color{red}r'_q},{\color{red}a'_{q+1}},r_q,\ldots,r_B\}\\ \rightarrow \mathcal{S} &= \{a_1,\ldots,a_q,\quad,\quad,r_q,\ldots,r_B\}.\end{aligned}$$

-

$$\begin{aligned}q\left[(\mathcal{M}',\theta')|(\mathcal{M},\theta)\right] &= p_{birth}\times\pi_q^A\times g(u_1,u_2)\\ &= p_{birth}\times\frac{r_q-a_q}{\sum_{b=1}^{B}(r_b-a_b)}\times 2!\frac{1}{(r_q-a_q-1)^2}.\end{aligned}$$

RJMCMC - Birth Move

Break down the acceptance probability for $(r'_q, a'_{q+1}]$:

$$\min\left\{1,\ \frac{p(\theta'|\boldsymbol{y},\mathcal{M}')q\left[(\mathcal{M},\theta)|(\mathcal{M}',\theta')\right]}{p(\theta|\boldsymbol{y},\mathcal{M})q\left[(\mathcal{M}',\theta')|(\mathcal{M},\theta)\right]}\left|\frac{\partial(\theta')}{\partial(\theta,u)}\right|\right\}.$$

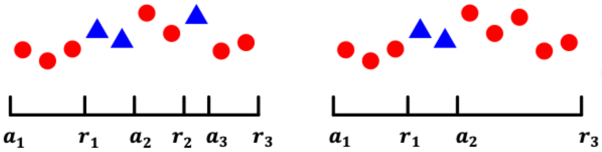
Jacobian:

$$\begin{aligned}\frac{\partial(\theta')}{\partial(\theta,u)} &= \frac{\partial(\mathcal{S}')}{\partial(\mathcal{S},u_1,u_2)}\\ &= \frac{\partial(r'_q, a'_{q+1})}{\partial(u_1,u_2)}\\ &= 1\end{aligned}$$

$\because u_1 = r'_q$ and $u_2 = a'_{q+1}$.

RJMCMC - Death Move

Remove an existing interval $(a_q, r_q]$ or $(r_q, a_{q+1}]$.



Death of a resting interval $(r_q, a_{q+1}]$:

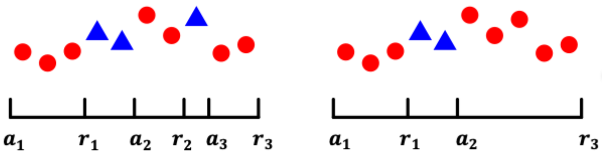
1) Choose $(r_q, a_{q+1}]$ to be removed.

$$\mathcal{S} = \{a_1, \ldots, a_q, r_q, a_{q+1}, r_{q+1}, \ldots, r_B\}$$
$$\mathcal{S}' = \{a_1, \ldots, a_q, \quad , \quad , r_{q+1}, \ldots, r_B\}$$

$$\text{dim}(\mathcal{S}') = \text{dim}(\mathcal{S}) - 2.$$

RJMCMC - Death Move

Remove an existing interval $(a_q, r_q]$ or $(r_q, a_{q+1}]$.



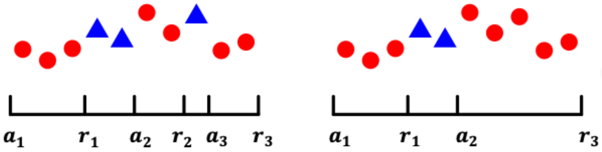
Involved probability:

$(r_q, a_{q+1}]$ is randomly chosen **inversely proportional** to its length:

$$\pi_q^R := \frac{L_q^R}{\sum_{b=1}^{B-1} L_b^R} \text{ where } L_b^R = \frac{1}{a_{b+1} - r_b}.$$

RJMCMC - Death Move

Remove an existing interval $(a_q, r_q]$ or $(r_q, a_{q+1}]$.



2) Auxiliary variables are not needed for the death move.

   – Remove the chosen $(r_q, a_{q+1}]$ from $\mathcal{S}$ can match the dimension between $\mathcal{S}'$ and $\mathcal{S}$.

$$\mathcal{S} = \{a_1, \ldots, a_q, \color{red}{r_q, a_{q+1}}, r_{q+1}, \ldots, r_B\}$$
$$\mathcal{S}' = \{a_1, \ldots, a_q, \quad \color{red}{,} \quad , r_{q+1}, \ldots, r_B\}$$

- The acceptance probability for the death move can be defined similarly to the birth move.
- $q\left[(\mathcal{M},\theta)|(\mathcal{M}',\theta')\right]$ from the birth move can now be defined:

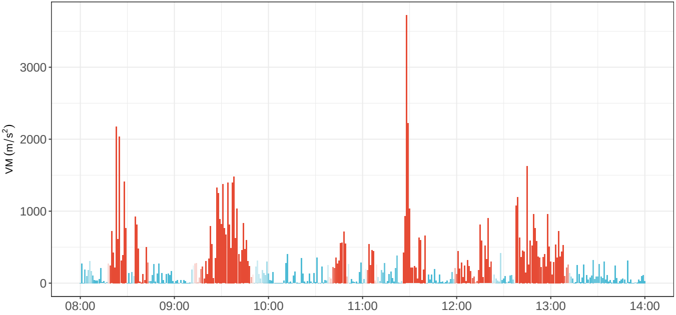
$$\begin{aligned}
\mathcal{S}' &= \{a_1, \ldots, a_q, r'_q, a'_{q+1}, r_q, \ldots, r_B\} \\
\rightarrow \mathcal{S} &= \{a_1, \ldots, a_q, \quad , \quad , r_q, \ldots, r_B\} \\
q\left[(\mathcal{M},\theta)|(\mathcal{M}',\theta')\right] &= p_{death} \times \pi_q^R \\
&= p_{death} \times \frac{L_q^R}{\sum_{b=1}^{B} L_b^R}, \\
&\text{where } L_q^R = \frac{1}{a'_{q+1} - r'_q}.
\end{aligned}$$

Content

Data

- A medical study of pediatric subjects with cerebral palsy.
  – recorded for 7 days.
- A total of 58 subjects with varying levels of movement ability.
- The most severely impaired case includes the group that cannot independently walk and use powered mobility.
- Three groups of subjects: ambulatory, marginally ambulatory, and non-ambulatory.

Results for an individual

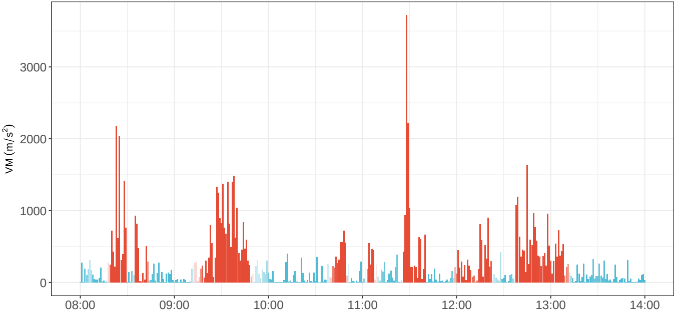


Posterior active state probability:

$$\hat{\rho}_t := M^{-1} \sum_{m=1}^{M} \sum_{b=1}^{\hat{B}_m} \mathbb{I}\left(y_t \in \hat{A}_{mb}\right) \quad \text{for } t = 1, \ldots, T.$$

$\hat{B}_m$: the estimated number of active intervals at m-th iteration.
$\hat{A}_{mb}$: b-th active interval at mth iteration.

Results for an individual



- **State uncertainty** by color-coding y_t in red if $\hat{\rho}_t \geq 0.5$ and blue otherwise.
- Unlike threshold-based approach using a single cut-off value.
- **Transparent** uncertainty measure, especially **near state boundaries** where uncertainty is higher.

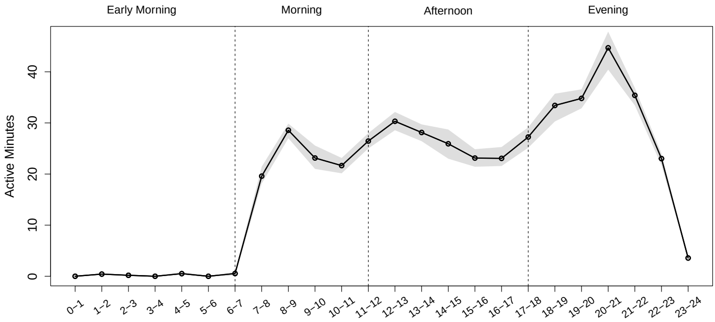
Results for an Individual

Summary of the five largest estimated mixing weights $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_K)$ and their corresponding Poisson parameters λ_k.

k		1	2	3	4	5
$\boldsymbol{\pi}$	Posterior mean	0.666	0.063	0.051	0.049	0.048
	95% CI	(0.652, 0.681)	(0.056, 0.069)	(0.045, 0.058)	(0.043, 0.055)	(0.043, 0.055)
$\boldsymbol{\lambda}$	Posterior mean	**0.005**	21.0	40.69	68.04	7.37
	95% CI	(0.003, 0.008)	(20.3, 21.8)	(39.5, 41.8)	(66.6, 69.6)	(7.0, 7.9)

- The first mixture component accounts for 66.6% of measurements labeled as the resting state.
- The remaining components spread out to capture occasional spikes.

Results for an individual - Daily Activity Pattern



- x-axis: 24 time periods.
- y-axis: proportion of the measurements labeled as active states.
- Black solid line: connecting posterior means.
- Shaded area: covering the 95% credible intervals at a given time.

[Summary of the measurements labeled as the active states]

$\hat{\mathcal{S}}_m := \left\{\hat{a}_{m1}, \hat{r}_{m1}, \ldots, \hat{a}_{m\hat{B}_m}, \hat{r}_{m\hat{B}_m}\right\}$: the estimated change points at m-th iteration.

1) **Active endurance**:

   – For $m = 1, \ldots, M$, $E_m := \sum_{b=1}^{\hat{B}_m} (\hat{r}_{mb} - \hat{a}_{mb}) / \hat{B}_m$, where M is the total number of MCMC after burn-in.

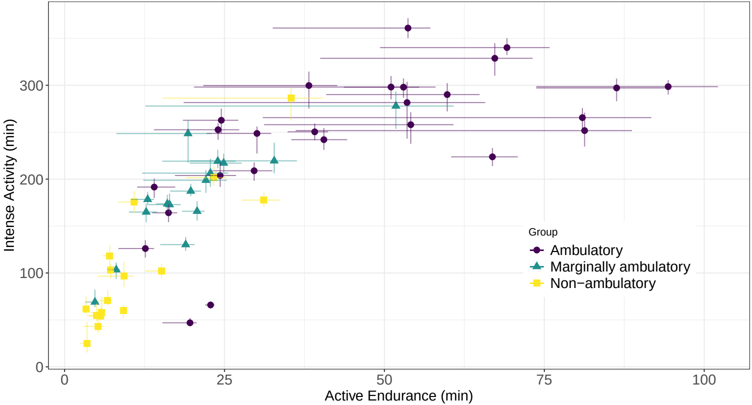
2) **Daily total duration of intense activity**:

   – A VM measurement is intense when $VM > \mu$.

   – For $m = 1, \ldots, M$, $I_m := \sum_{t: y_t \in \boldsymbol{y}_a} \mathbb{I}(y_t > \hat{\mu}_m)/7$, where $\hat{\mu}_m = m$-th estimated mean of $f_A \sim N(\mu\mathbf{1}, \sigma^2\boldsymbol{\Sigma})$.

Group Level Comparison

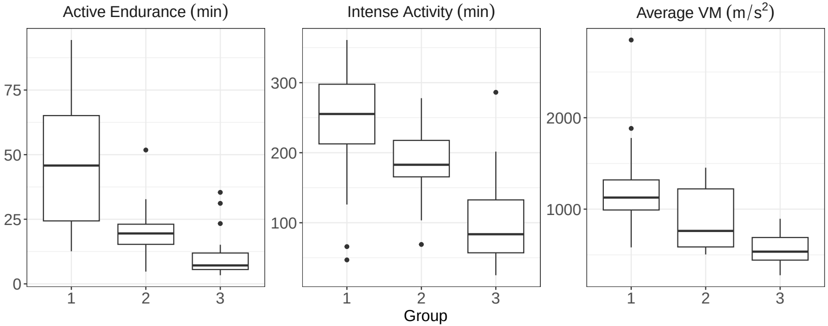
Posterior means and 95% credible intervals of active endurance and intense activity for each subject in minutes by group.



Severity of physical challenges is strongly associated with the amount of activities.

Group Level Comparison

Box plots treating each individual posterior mean as a data point.



- Average VM: simple mean of the measurements labeled as active states.
- First two suggested metrics: better to represent the differences between the groups & straightforward interpretability.

Content

Conclusions

Suggested RJMCMC algorithm

- enables uncertainty quantification related to state labeling.
- is preferred over heuristic algorithms due to its data-driven inference.
- allows for the calculation of various metrics beyond simple summary statistics.
- provides practitioners with more flexibility to process raw datasets tailored to their research.

Limitations and Future Works

- Active states include all types of activity: exercising, walking, and running, etc.

  $\rightarrow$ Develop more elaborate models that can capture the difference between activities.
- Several tools to prevent abrupt changes from labeling as a different state are not enough.

  $\rightarrow$ Filter noise more precisely.
- Current MCMC takes a long time to run.

  $\rightarrow$ Recently, deep learning models are studied, but still requiring training sets.

  $\rightarrow$ Developing unsupervised deep learning models for wearable sensor measurements.

Thank you !

Green, P. J. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika 82*(4), 711–732.

Ishwaran, H. and L. F. James (2001). Gibbs sampling methods for stick-breaking priors. *Journal of the American Statistical Association 96*(453), 161–173.

Lafferty, J. D., A. McCallum, and F. C. N. Pereira (2001). Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In C. E. Brodley and A. P. Danyluk (Eds.), *Proceedings of the Eighteenth International Conference on Machine Learning (ICML 2001), Williams College, Williamstown, MA, USA, June 28 - July 1, 2001*, pp. 282–289. Morgan Kaufmann.

Neal, R. M. (1999, 08). Regression and Classification Using Gaussian Process Priors. In *Bayesian Statistics 6: Proceedings of the Sixth Valencia International Meeting June 6-10, 1998*. Oxford University Press.