Web-based Supplementary Materials for “Movement Prediction Using Accelerometers in a Human Population”

by

Luo Xiao, Bing He, Annemarie Koster, Paolo Caserotti, Brittney Lange-Maia, Nancy W. Glynn, Tamara B. Harris and Ciprian M. Crainiceanu

A Additional Figures

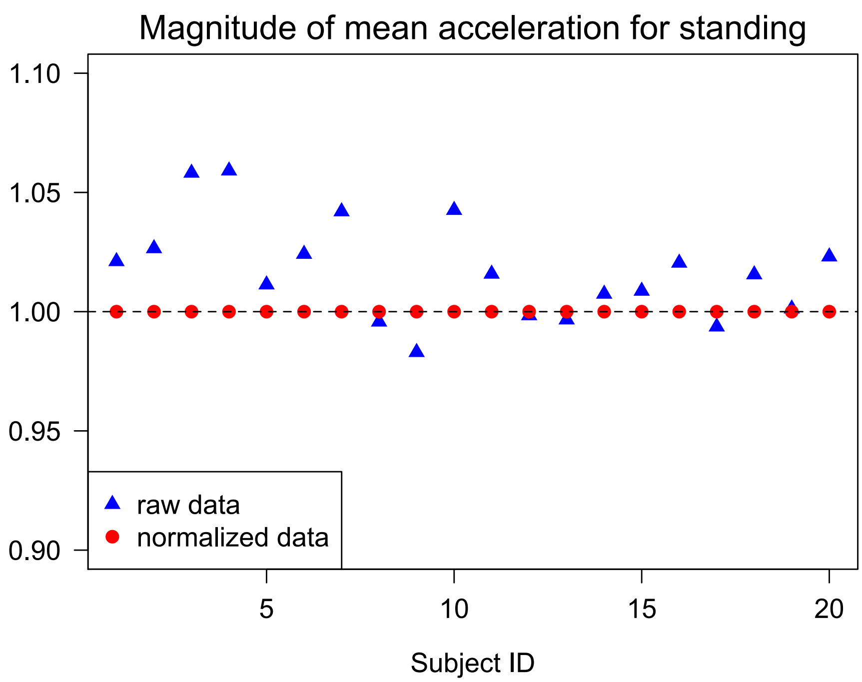

Figure Web-1: Magnitude of the mean acceleration vector during standing (expressed in g units) using the raw and normalized data. Blue triangles are for raw data while red circles are for normalized data.

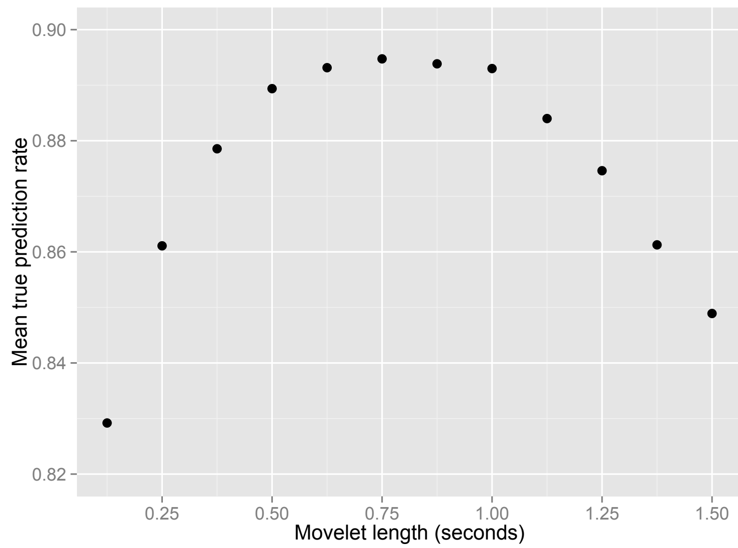

Figure Web-2: Cross-validated mean true prediction rate using different movelet lengths for the training data.

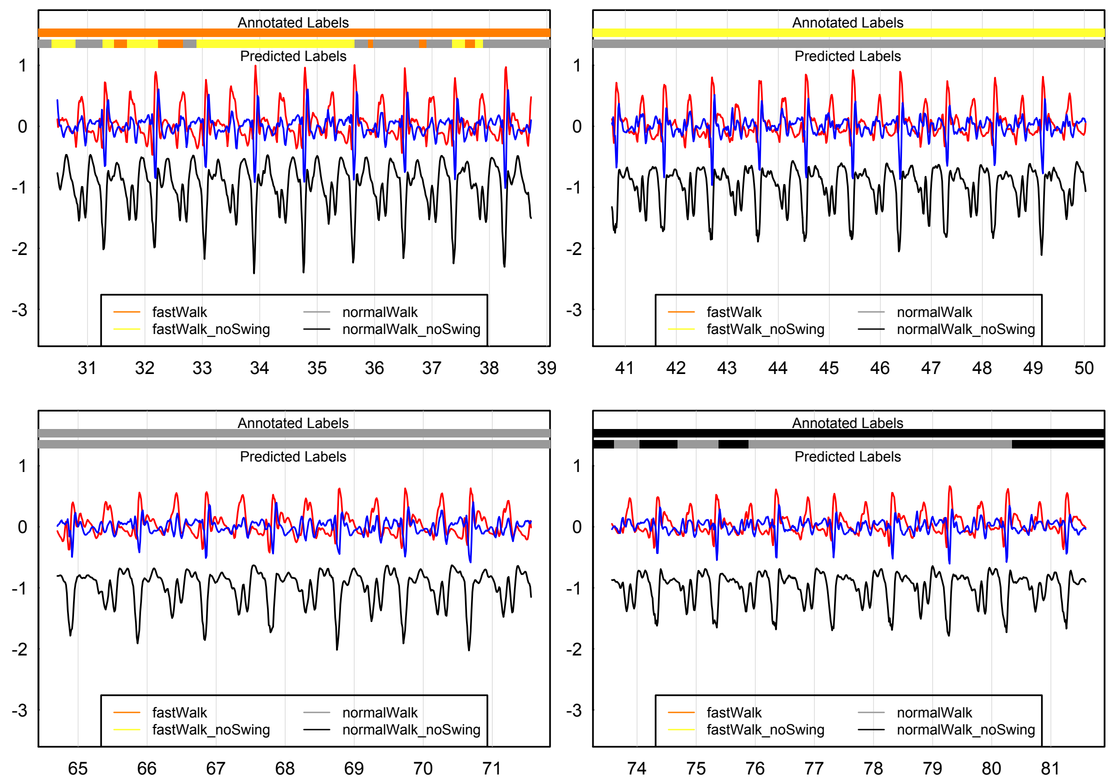

Figure Web-3: Prediction results of one subject's four types of walking. The top panels display data for normalWalk (left column) and normalWalk_noSwing (right column), the bottom panels display data for fastWalk (left column) and fastWalking_noSwing (right column). The activity types can also be distinguished by the annotated labels in each plot. For the accelerometry data, the black lines are for the up-down axis, the red lines are for the forward-backward axis, and the blue lines are for the left-right axis.

B A Test for Systematic Bias

Let $\mathbf{X} \in \mathbb{R}^3$ be the acceleration vector at an observation point when the subject is standing still. Suppose that $\mathbf{X}$ follows a multivariate normal distribution with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$. Let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ be i.i.d. copies of $\mathbf{X}$. Then $\bar{\mathbf{X}} = n^{-1} \sum_{i=1}^n \mathbf{X}_i$ is normal with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}/n$. Testing if there is systematic bias in the observations is to testing $\|\boldsymbol{\mu}\| = 1$ where $\|\cdot\|$ denotes the Euclidean norm. We consider the testing statistic $\|\bar{\mathbf{X}}\|^2$, which has mean $\|\boldsymbol{\mu}\|^2 + tr(\boldsymbol{\Sigma})/n$ and variance $\mathrm{var}(\|\bar{\boldsymbol{X}}\|^2)$. Here $tr(\cdot)$ denotes the trace of a square matrix, i.e., the sum of the diagonal entries. The derivation of the variance term is more involved. We let $\mathbf{ODO}^T$ be the eigendecomposition of $\boldsymbol{\Sigma}$, where $\mathbf{O}$ is an orthogonal matrix with $\mathbf{O}^T\mathbf{O} = \mathbf{O}\mathbf{O}^T = \mathbf{I}_3$ and $\mathbf{D}$ is a diagonal matrix with the diagonal entries d_1, d_2 and d_3. Now let $\mathbf{Y} = (Y_1, Y_2, Y_3)^T = \mathbf{O}^T\bar{\mathbf{X}}$, then $\mathbf{Y}$ is normal with mean $\boldsymbol{\mu}_y = (\mu_{y1}, \mu_{y2}, \mu_{y3})^T = \mathbf{O}^T\boldsymbol{\mu}$ and covariance matrix $\mathbf{D}/n$. It is easy to show that

$$\mathrm{var}(\|\bar{\mathbf{X}}\|^2) = \mathrm{var}(\|\mathbf{Y}\|^2) = \sum_{k=1}^3 \mathrm{var}(Y_k^2) = \sum_{k=1}^3 (6\mu_{yk}^2 d_k/n + 3d_k^2/n^2).$$

Hence

$$\begin{aligned}
\mathrm{var}(\|\bar{\mathbf{X}}\|^2) &= \sum_{k=1}^3 (6\mu_{yk}^2 d_k/n + 3d_k^2/n^2) \\
&= \frac{6}{n}\boldsymbol{\mu}_y^T \mathbf{D}\boldsymbol{\mu}_y + \frac{3}{n^2} tr(\boldsymbol{\Sigma}^2) \\
&= \frac{6}{n}\boldsymbol{\mu}^T \mathbf{ODO}'\boldsymbol{\mu} + \frac{3}{n^2} tr(\boldsymbol{\Sigma}^2) \\
&= \frac{6}{n}\boldsymbol{\mu}^T \boldsymbol{\Sigma}\boldsymbol{\mu} + \frac{3}{n^2} tr(\boldsymbol{\Sigma}^2).
\end{aligned}$$

By the central limit theorem, $\frac{\|\bar{\mathbf{X}}\|^2 - \|\boldsymbol{\mu}\|^2 - tr(\boldsymbol{\Sigma})/n}{\sqrt{\mathrm{var}(\|\bar{\mathbf{X}}\|^2)}}$ is approximately normal. Then an α-level rejection region for testing $\|\boldsymbol{\mu}\| = 1$ is given by

$$\left|\|\bar{\mathbf{X}}\|^2 - 1\right| > z_{\alpha/2}\sqrt{\frac{6}{n}\boldsymbol{\mu}^T\boldsymbol{\Sigma}\boldsymbol{\mu}}$$

Note that we dropped the term $tr(\boldsymbol{\Sigma})/n$ in the numerator and the term $3tr(\boldsymbol{\Sigma}^2)/n^2$ in the denominator as they are of smaller order than $\|\boldsymbol{\mu}\|^2$ and $6\boldsymbol{\mu}^T\boldsymbol{\Sigma}\boldsymbol{\mu}/n$, respectively. The term $\boldsymbol{\mu}^T\boldsymbol{\Sigma}\boldsymbol{\mu}$ is unknown and needs to be estimated. Since under the null hypothesis that $\|\boldsymbol{\mu}\| = 1$ we can derive $\boldsymbol{\mu}^T\boldsymbol{\Sigma}\boldsymbol{\mu} \leq \|\boldsymbol{\Sigma}\|_{op}$, where $\|\cdot\|_{op}$ is the operator norm of a matrix, we use instead a conservative rejection region

$$\left|\|\bar{\mathbf{X}}\|^2 - 1\right| > z_{\alpha/2}\sqrt{\frac{6}{n}\|\hat{\boldsymbol{\Sigma}}\|_{op}},$$

Table Web-1: Testing statistic T for the 20 subjects

Subject	T
1	209.38
2	76.11
3	473.84
4	103.94
5	71.17
6	209.28
7	183.57
8	84.22
9	134.98
10	365.15
11	228.35
12	0.65
13	1.93
14	22.47
15	9.59
16	191.47
17	83.44
18	11.68
19	11.91
20	165.02

where $\hat{\mathbf{\Sigma}}$ is the sample covariance matrix from the sample $\{\mathbf{X}_1, \ldots, \mathbf{X}_n\}$. We use $\alpha = 0.05$ so that $z_{\alpha/2} = 1.96$. We display the value of the term

$$T = \frac{\left|\|\bar{\mathbf{X}}\|^2 - 1\right|}{\sqrt{\frac{6}{n}\|\hat{\mathbf{\Sigma}}\|_{op}}}$$

for all subjects in Table Web-1. The results show that except for subjects 12 and 13, the null hypothesis of $\|\boldsymbol{\mu}\|$ is always rejected.

C Derivation of Rotation Matrices

Let $\mathbf{a}_1$ and $\mathbf{a}_2$ be two vectors in $\mathbb{R}^3$ and $\mathbf{a}_1 \times \mathbf{a}_2 \neq 0$. Let

$$\mathbf{b}_2 = \frac{\mathbf{a}_2}{\|\mathbf{a}_2\|_2},$$
$$\mathbf{b}_1 = \frac{\mathbf{a}_1 - (\mathbf{a}_1^T\mathbf{b}_2)\mathbf{b}_2}{\|\mathbf{a}_1 - (\mathbf{a}_1^T\mathbf{b}_2)\mathbf{b}_2\|_2}.$$

Then $\mathbf{b}_1$ and $\mathbf{b}_2$ are two unit vectors and are orthogonal to each other. We write $\mathbf{a}_1$ and $\mathbf{a}_2$ as

$$\begin{aligned}\mathbf{a}_2 &= c_2\mathbf{b}_2,\\ \mathbf{a}_1 &= c_1\mathbf{b}_1 + c_3\mathbf{b}_2.\end{aligned}$$

Then $c_1 > 0, c_2 > 0$ and $c_3 > 0$.

Lemma 1. *Let*

$$\mathbf{R}^* = \arg \min_{\mathbf{R}^T=\mathbf{R}^{-1} \text{ and } \mathbf{e}_3^T\mathbf{R}(\mathbf{a}_1\times\mathbf{a}_2)>0} \left(\|\mathbf{R}\mathbf{a}_1 + \mathbf{e}_1\|_2^2 + \|\mathbf{R}\mathbf{a}_2 + \mathbf{e}_2\|_2^2\right).$$

Then $\mathbf{R}^*$ *is unique with the expression*

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0\\ \sin(\theta) & \cos(\theta) & 0\\ 0 & 0 & 1\end{pmatrix}^T [\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_1\times\mathbf{b}_2]^T,$$

where

$$\begin{aligned}\cos(\theta) &= -\frac{c_1+c_3}{\sqrt{(c_1+c_2)^2+c_3^2}},\\ \sin(\theta) &= -\frac{c_2}{\sqrt{(c_1+c_2)^2+c_3^2}}.\end{aligned}$$

Proof. For an arbitrary rotation matrix $\mathbf{R}$, $\mathbf{R}^T = \mathbf{R}^{-1}$ is also a rotation matrix. Hence $\mathbf{R}^T\mathbf{e}_1$ and $\mathbf{R}^T\mathbf{e}_1$ remain orthogonal unit vectors. For the minimization problem, there exists an $\theta \in [0, \pi]$ such that

$$\begin{aligned}\mathbf{R}^T\mathbf{e}_1 &= \cos(\theta)\mathbf{b}_1 + \sin(\theta)\mathbf{b}_2,\\ \mathbf{R}^T\mathbf{e}_2 &= -\sin(\theta)\mathbf{b}_1 + \cos(\theta)\mathbf{b}_2.\end{aligned} \tag{A-1}$$

It follows that

$$\begin{aligned}&\|\mathbf{R}\mathbf{a}_1 + \mathbf{e}_1\|_2^2 + \|\mathbf{R}\mathbf{a}_2 + \mathbf{e}_2\|_2^2\\ &= \|\mathbf{a}_1 + \mathbf{R}^T\mathbf{e}_1\|_2^2 + \|\mathbf{a}_2 + \mathbf{R}^T\mathbf{e}_2\|_2^2\\ &= \|(c_1+\cos(\theta))\mathbf{b}_1 + (c_3+\sin(\theta))\mathbf{b}_2\|_2^2 + \|-\sin(\theta)\mathbf{b}_1 + (c_2+\cos(\theta))\mathbf{b}_2\|_2^2\\ &= (c_1+\cos(\theta))^2 + (c_3+\sin(\theta))^2 + \sin(\theta)^2 + (c_2+\cos(\theta))^2\\ &= 2 + c_1^2 + c_2^2 + c_3^2 + 2(c_1+c_2)\cos(\theta) + 2c_3\sin(\theta).\end{aligned}$$

Therefore, $\|\mathbf{R}\mathbf{a}_1 + \mathbf{e}_1\|_2^2 + \|\mathbf{R}\mathbf{a}_2 + \mathbf{e}_2\|_2^2$ is minimized if $\cos(\theta) = -(c_1+c_2)/\sqrt{(c_1+c_2)^2+c_3^2}$ and $\sin(\theta) = -c_3/\sqrt{(c_1+c_2)^2+c_3^2}$. By (A-1),

$$\begin{aligned}\mathbf{R}^T\mathbf{e}_3 &= \mathbf{R}^T(\mathbf{e}_1\times\mathbf{e}_2)\\ &= (\mathbf{R}^T\mathbf{e}_1)\times(\mathbf{R}^T\mathbf{e}_2)\\ &= \mathbf{b}_1\times\mathbf{b}_2.\end{aligned}$$

Then

$$\mathbf{R}^T[\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3] = [\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_1 \times \mathbf{b}_2] \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

It's easy to verify that $\mathbf{R}^{-1} = \mathbf{R}^T$ and the proof is complete. □

D An Alternative Approach for Normalizing Accelerometry Data

We propose a method for jointly estimating the rotation matrix $\mathbf{R}$ and the bias vector $\mathbf{b}$ for standing. We minimize the following objective function

$$\min_{\mathbf{R}^T=\mathbf{R}^{-1}, \mathbf{b}} \|\mathbf{R}\mathbf{a}_1 + \mathbf{e}_1 - \mathbf{b}\|_2^2 + \|\mathbf{R}\mathbf{a}_2 + \mathbf{e}_2\|_2^2.$$

It can be shown easily that the minimizers to the above are $\hat{\mathbf{b}} = \hat{\mathbf{R}}\mathbf{a}_1 + \mathbf{e}_1$ and

$$\hat{\mathbf{R}} = \arg \min_{\mathbf{R}^T=\mathbf{R}^{-1}} \|\mathbf{R}\mathbf{a}_2 + \mathbf{e}_2\|_2^2.$$

However, $\hat{\mathbf{R}}$ is not unique and satisfies $\hat{\mathbf{R}}^T = \hat{\mathbf{R}}^{-1}$ and $\hat{\mathbf{R}}^T\mathbf{e}_2 = -\mathbf{b}_2$. Here we are using the notation at the beginning of last section. The idea is that we expect the bias $\mathbf{b}$ to be small and hence we select the final estimate of rotation matrix by

$$\mathbf{R}^* = \arg \min_{\hat{\mathbf{R}}^T=\hat{\mathbf{R}}^{-1}, \hat{\mathbf{R}}^T\mathbf{e}_2=-\mathbf{b}_2} \|\mathbf{R}\mathbf{a}_1 + \mathbf{e}_1\|_2^2.$$

It can be shown that $\mathbf{R}^{*,T}\mathbf{e}_1 = -\mathbf{b}_1$. Therefore with the conditions

$$\mathbf{R}^{*,T}\mathbf{e}_1 = -\mathbf{b}_1, \mathbf{R}^{*,T}\mathbf{e}_2 = -\mathbf{b}_2$$

and that $\mathbf{e}_3^T\mathbf{R}^*(\mathbf{a}_1 \times \mathbf{a}_2) > 0$, we obtain

$$\mathbf{R}^* = [-\mathbf{b}_1, -\mathbf{b}_2, \mathbf{b}_1 \times \mathbf{b}_2]^T.$$

The results from the method proposed in the paper and the above alternative method can be compared. It is easy to see that if c_3 is zero, i.e., $\mathbf{a}_1$ and $\mathbf{a}_2$ are orthogonal, expected if no bias in the measurement, then $\cos(\theta) = -1$ and we get the same results. If c_3 is far from zero, then we will get different results. For the data example, we found c_3 small for all subjects and hence both methods give very similar results.