『入門 機械学習による異常検知 — R による実践ガイド —』（コロナ社、2015）初版第 1 刷の正誤表

Tsuyoshi Idé (井手 剛) ide@ide-research.net

平成 28 年 9 月 13 日

1 異常検知の基本的な考え方

- p.11
  - 誤: 例えば,「>」の右に,56*log(10^ (-20)) と打つと,$50 \ln 10^{20}$ の答えが出てきます。
  - 正: 例えば,「>」の右に,56*log(10^ (-20)) と打つと,$56 \ln 10^{-20}$ の答えが出てきます。

2 正規分布に従うデータからの異常検知

- p.28, 式 (2.17)。
  - 誤: $0 = \dfrac{\partial L}{\partial \mu} = -\dfrac{1}{2\sigma^2}\sum_{n=1}^{N}(\mu - x^{(n)})$
  - 正: $0 = \dfrac{\partial L}{\partial \mu} = -\dfrac{1}{\sigma^2}\sum_{n=1}^{N}(\mu - x^{(n)})$
- p.56, 式 (2.57)
  - 誤 $\mathcal{S}(t \mid N-1, 0) = \dfrac{\Gamma[(m+1)/2]}{\Gamma(m/2)}\sqrt{\dfrac{1}{\pi m}}\left[1 + \dfrac{t^2}{m}\right]^{-(m+1)/2}$
  - 正 $\mathcal{S}(t \mid m, 0) = \dfrac{\Gamma[(m+1)/2]}{\Gamma(m/2)}\sqrt{\dfrac{1}{\pi m}}\left[1 + \dfrac{t^2}{m}\right]^{-(m+1)/2}$
- 章末問題の最後、【7】での表現がやや不正確でした。
  - 誤: 各変数が統計的に独立なとき
  - 正: 各変数が統計的に無相関であるとき

3 非正規データからの異常検知

- p.71、上から 3 行目（森川浩司様のご指摘に感謝いたします）。
  - 誤: 標準偏差 `sig0=3` の正規分布
  - 正: 標準偏差 `sig1=3` の正規分布
- p.83、実行例 3.6 のすぐ上（森川浩司様のご指摘に感謝いたします）。
  - 誤: 評価します（実行例 3.6）。
  - 正: 評価します。次の実行例 3.6 は、標本数を `n`，カーネル行列を `K` として格納した前提の計算例です。
- p.92、実行例 3.8、1 行目（山下智輝様のご指摘に感謝いたします）。
  - 誤: `$` 混合比を取り出す
  - 正: `#` 混合比を取り出す

4 性能評価の方法

- p.112、上から 3 行目（佐々木俊久様のご指摘に感謝いたします)。
  - 誤: ソートされた異常度 anomaly_sorted の関数として
  - 正: ソートされた異常度 score_sorted の関数として

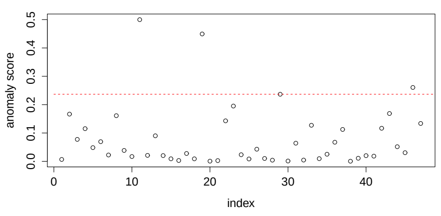
5 不要な次元を含むデータからの異常検知

- p.126 最初の式。森川浩司様のご指摘に感謝いたします。
  - 誤: $\dfrac{1}{N}\sum_{i=1}^{N}\boldsymbol{u}^\top\boldsymbol{x}^{(n)} = \dfrac{1}{N}\boldsymbol{u}^\top\sum_{n=1}^{N}\boldsymbol{x}^{(n)} = \dfrac{1}{N}\boldsymbol{u}^\top\hat{\boldsymbol{\mu}}$
  - 正: $\dfrac{1}{N}\sum_{n=1}^{N}\boldsymbol{u}^\top\boldsymbol{x}^{(n)} = \dfrac{1}{N}\boldsymbol{u}^\top\sum_{n=1}^{N}\boldsymbol{x}^{(n)} = \boldsymbol{u}^\top\hat{\boldsymbol{\mu}}$
- p.130、式 (5.18) の 2 行下。森川浩司様のご指摘に感謝いたします。
  - 誤: 正規直交ベクトルを $r - M$ 本加えて
  - 正: 正規直交ベクトルを $M - r$ 本加えて
- p.130、下から 4 行目。池田弘様のご指摘に感謝いたします。
  - 誤: 式 (5.21) から、Λ_M のゼロ対角要素に対応する部分を省略して
  - 正: 式 (5.20) から、Λ_M のゼロ対角要素に対応する部分を省略して
- p.132 上から 4 行目。吉川岳様のご指摘に感謝いたします。
  - 誤: $N \times N$ の中心化グラム行列の上位 r 個の規格化された固有ベクトル $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m$ を求め
  - 正: $N \times N$ の中心化グラム行列の上位 m 個の規格化された固有ベクトル $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m$ を求め
- p.132、下から 3 行目。森川浩司様のご指摘に感謝いたします。
  - 誤: M 次元空間の正規直交基底 r 個が
  - 正: M 次元空間の正規直交基底 m 個が
- p.144、式 (5.47) の 3 行目。森川浩司様のご指摘に感謝いたします。
  - 誤: $-\dfrac{M}{2}\ln(2\pi\sigma^2) +$ 定数

  - 正: $-\dfrac{MN}{2}\ln(2\pi\sigma^2) + \text{定数}$
- p.144、式 (5.49)。那須翔太様のご指摘に感謝いたします。
  - 誤: $\langle \boldsymbol{z}^{(n)} \boldsymbol{z}^{(n)\top} \rangle = \sigma^2 \mathsf{M}'^{-1} + \langle \boldsymbol{z}^{(n)} \rangle \langle \boldsymbol{z}^{(n)} \rangle^\top$
  - 正: $\langle \boldsymbol{z}^{(n)} \boldsymbol{z}^{(n)\top} \rangle = \sigma^{2\prime} \mathsf{M}'^{-1} + \langle \boldsymbol{z}^{(n)} \rangle \langle \boldsymbol{z}^{(n)} \rangle^\top$
- 章末問題の【5】での表現がやや不正確でした。
  - 誤: どの要素もゼロではない任意の M 次元単位ベクトル $\boldsymbol{z}$ に対して、実対称行列 A を使い
  - 正:実対称行列 A が重複のない最大固有値を持つとし、対応する固有ベクトルを $\boldsymbol{u}_1$ とします。$\boldsymbol{u}_1^\top \boldsymbol{z} \neq 0$ を満たす任意の M 次元単位ベクトル $\boldsymbol{z}$ を考え

6 入力と出力があるデータからの異常検知

- p.168、最初の式。森川浩司様のご指摘に感謝いたします。
  - 誤: $[\cdots]^{-1} = \mathsf{A}^{-1} + \mathsf{A}^{-1}\tilde{\boldsymbol{x}}^{(n)}[1 - \tilde{\boldsymbol{x}}^{(n)\top}\mathsf{A}\tilde{\boldsymbol{x}}^{(n)}]\tilde{\boldsymbol{x}}^{(n)\top}\mathsf{A}^{-1}$
  - 正: $[\cdots]^{-1} = \mathsf{A}^{-1} + \mathsf{A}^{-1}\tilde{\boldsymbol{x}}^{(n)}[1 - \tilde{\boldsymbol{x}}^{(n)\top}\mathsf{A}^{-1}\tilde{\boldsymbol{x}}^{(n)}]^{-1}\tilde{\boldsymbol{x}}^{(n)\top}\mathsf{A}^{-1}$
- p.168、式 (6.23)。
  - 誤： $e(\lambda) = \dfrac{1}{N}\displaystyle\sum_{N=1}^{N} \dfrac{1}{(1-\mathsf{H}_{n,n})^2}\left[\tilde{y}^{(n)} - \tilde{\boldsymbol{x}}^{(n)\top}\hat{\boldsymbol{\alpha}}_{\text{ridge}}\right]^2$
  - 正： $e(\lambda) = \dfrac{1}{N}\displaystyle\sum_{n=1}^{N} \dfrac{1}{(1-\mathsf{H}_{n,n})^2}\left[\tilde{y}^{(n)} - \tilde{\boldsymbol{x}}^{(n)\top}\hat{\boldsymbol{\alpha}}_{\text{ridge}}\right]^2$
- p.172、図 6.3 (b)。これは出版社との連携ミスで、別の図が使われてしまったようです。原図は下記です。森川浩司様のご指摘に感謝いたします。



- p.172、実行例 6.2。5 行目。異常度の計算にミスがありました。佐々木俊久様のご指摘に感謝いたします。
  - 誤: `a <- (as.numeric(ypred) - y)^2/((1 - TrHN)*sig2)`
  - 正: `a <- (as.numeric(ypred) - y)^2/((1 - TrHN)^2*sig2)`
- p.174、真ん中の式の上。那須翔太様のご指摘に感謝いたします。

 - 誤: 生の $\boldsymbol{x}$ を扱う代わりに、 個の正規直交基底
 - 正: 生の $\boldsymbol{x}$ を扱う代わりに、m 個の正規直交基底
- p.176、下から 5 行目。那須翔太様のご指摘に感謝いたします。
 - 誤: 基底 $\boldsymbol{p}$ を $[c_1, \ldots, c_M]^\top$ のように求めます。
 - 正: 基底 $\boldsymbol{p}_1$ を $[c_1, \ldots, c_M]^\top$ のように求めます。
- p.184、上から 8 行目。大堀龍一様のご指摘に感謝いたします。
 - 誤: そのようにして特異ベクトルを $(\tilde{\boldsymbol{\alpha}}^1, \tilde{\boldsymbol{\beta}}^1), (\tilde{\boldsymbol{\alpha}}^2 \tilde{\boldsymbol{\beta}}^2), \ldots$ のように求めたとしましょう。
 - 正: そのようにして特異ベクトルを $(\tilde{\boldsymbol{\alpha}}^1, \tilde{\boldsymbol{\beta}}^1), (\tilde{\boldsymbol{\alpha}}^2, \tilde{\boldsymbol{\beta}}^2), \ldots$ のように求めたとしましょう。
- p.189、3 番目の式。森川浩司様のご指摘に感謝いたします。
 - 誤: $\left(\sigma^2 \mathsf{I}_N + \frac{\sigma^2}{\hat{\lambda}} \Phi^\top \Phi\right)^{-1} = \frac{1}{\sigma^2}\left(\mathsf{I}_{M_\phi} - \Phi^\top \mathsf{A}^{-1} \Phi\right)$
 - 正: $\left(\sigma^2 \mathsf{I}_N + \frac{\sigma^2}{\hat{\lambda}} \Phi^\top \Phi\right)^{-1} = \frac{1}{\sigma^2}\left(\mathsf{I}_N - \Phi^\top \mathsf{A}^{-1} \Phi\right)$
- p.189、式 (6.66) の 2 行目。森川浩司様のご指摘に感謝いたします。
 - 誤: $-\frac{1}{2\sigma^2} \boldsymbol{y}_N^\top \left(\mathsf{I}_{M_\phi} - \Phi^\top \mathsf{A}^{-1} \Phi\right) \boldsymbol{y}_N$
 - 正: $-\frac{1}{2\sigma^2} \boldsymbol{y}_N^\top \left(\mathsf{I}_N - \Phi^\top \mathsf{A}^{-1} \Phi\right) \boldsymbol{y}_N$

7 時系列データの異常検知

- p.198 実行例 7.1 の 2 行目。森川浩司様のご指摘に感謝いたします。
  - 誤:
    ```
    dt <- read.table(file=paste(dir,"qtdbsel102.txt",sep=""))
    ```
  - 正:
    ```
    X <- read.table(file=paste(dir,"qtdbsel102.txt",sep=""))
    ```
- p.205 式 (7.10) 。大堀龍一様のご指摘に感謝いたします。
  - 誤: $\xi^{(t)} \approx \alpha_1 \xi^{(t-1)} + \alpha_2 \xi^{(t-2)} + \cdots + \alpha_r \xi^{(t-r)}$
  - 正: $\xi^{(t)} \approx \alpha_1 \xi^{(t-r)} + \alpha_2 \xi^{(t-r+1)} + \cdots + \alpha_{r-1} \xi^{(t-2)} + \alpha_r \xi^{(t-1)}$
- p.207 式 (7.15) 。大堀龍一様のご指摘に感謝いたします。
  - 誤: $\boldsymbol{\xi}^{(t)} \approx \mathsf{A}_1 \boldsymbol{\xi}^{(t-1)} + \mathsf{A}_2 \boldsymbol{\xi}^{(t-2)} + \cdots + \mathsf{A}_r \boldsymbol{\xi}^{(t-r)}$
  - 正: $\boldsymbol{\xi}^{(t)} \approx \mathsf{A}_1 \boldsymbol{\xi}^{(t-r)} + \mathsf{A}_2 \boldsymbol{\xi}^{(t-r+1)} + \cdots + \mathsf{A}_{r-1} \boldsymbol{\xi}^{(t-2)} + \mathsf{A}_r \boldsymbol{\xi}^{(t-1)}$
- p.215 図 7.9 のキャプション。小野剛様のご指摘に感謝いたします。
  - 誤: 図 7.9 部分空間同定法の説明。
  - 正: 図 7.9 部分空間同定法の説明。本文中の設定だと p と f には一般に多少の重なりが生じます（図 7.4 も参照)。

8 よくある悩みとその対処法

現時点で判明している誤りはありません。

付　　　録

- 定理 A.6 の式 (A.38) の右辺で余計な負号がひとつありました。何度も使う公式なので罪は重いです。申し訳ありません...。
  - 誤:
  $$\int_{-\infty}^{+\infty} \mathrm{d}x \ \exp(-ax^2+bx+c) = \sqrt{\frac{\pi}{a}} \exp\left(-\frac{b^2}{4a}+c\right)$$
  - 正:
  $$\int_{-\infty}^{+\infty} \mathrm{d}x \ \exp(-ax^2+bx+c) = \sqrt{\frac{\pi}{a}} \exp\left(\frac{b^2}{4a}+c\right)$$
- p.258、下から 2 番目の式。森川浩司様のご指摘に感謝いたします。
  - 誤: $\Lambda_{aa}^{-1} = \Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba}$ および $\Lambda_{aa}^{-1}\Lambda_{ab} = -\Sigma_{ab}^{-1}\Sigma_{bb}$
  - 正: $\Lambda_{aa}^{-1} = \Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba}$ および $\Lambda_{aa}^{-1}\Lambda_{ab} = -\Sigma_{ab}\Sigma_{bb}^{-1}$
- p.259、定理 A.9。森川浩司様のご指摘に感謝いたします。
  - 誤: $p(\boldsymbol{x} \mid \boldsymbol{x})$ および $p(\boldsymbol{y})$ は次で与えられる。
  - 正: $p(\boldsymbol{x} \mid \boldsymbol{y})$ および $p(\boldsymbol{y})$ は次で与えられる。
- p.264、(A.65) と (A.66) の間の式。別に間違いではないのですが、こっちの方がわかりやすいかと思うので書き換えます。
  - 古:
  $$f(\boldsymbol{x}+\boldsymbol{\epsilon}) \simeq f(\boldsymbol{x}) + \boldsymbol{\epsilon}^\top \nabla f(\boldsymbol{x}) \quad \text{すなわち} \quad f(\boldsymbol{x}) + \boldsymbol{\epsilon}^\top \frac{\partial f(\boldsymbol{x})}{\partial \boldsymbol{x}}$$
  - 新（当初 ∇ が抜けていたのを 2015/05/06 に修正）:
  $$f(\boldsymbol{x}+\boldsymbol{\epsilon}) \simeq f(\boldsymbol{x})+\boldsymbol{\epsilon}^\top \nabla f(\boldsymbol{x}) \quad \text{および} \quad g(\boldsymbol{x}+\boldsymbol{\epsilon}) \simeq g(\boldsymbol{x})+\boldsymbol{\epsilon}^\top \nabla g(\boldsymbol{x})$$
- p.265、下から 4 行目。$g(\boldsymbol{x}) = c$ の形の制約についての記述が乱暴でしたので修正いたします。森川浩司様のご指摘に感謝いたします。

– 誤: $g(\boldsymbol{x}) = c$ という制約を加えても，同様に定理が成立することは明らかです。
– 正: $g(\boldsymbol{x}) = c$ という制約を加えても，式 (A.68) が成立することは明らかです。

索　　引

- 非正値二次形式の訳語に誤りがありました。なお、negative semidefinite quadratic form という言い方もあります。
  - 誤: non-positive quadrature
  - 正: non-positive quadratic form