Прогнозирование намерений по сигналам мозга ECoG*

Калиниченко О. И., Ремизова А. С.
[1] Московский физико-технический институт

Работа посвящена построению системы тестирования прогностических моделей. Рассматривается случай коррелированных входных сигналов высокой размерности. В качестве прикладной задачи рассматривается задача предсказания намерений по сигналам головного мозга. Входные данные – сигналы электрокортикограммы (ECoG). Для выявления и устранения скрытых зависимостей используются методы снижения размерности пространства и отбора признаков. Предложенная система тестирования оценивает качество прогноза моделей и проводит анализ ошибки. Вычислительный эксперимент проводится на реальных данных ECoG.

Ключевые слова: *декодирование временных рядов, PLS, QPFS.*

Введение

Работа посвящена исследованию методов моделирования нейросетевого интерфейса (BCI) [4]. Входные данные – сигналы мозга, полученные с помощью электрокортикографии (ECoG) и электроэнцефалографии (EEG). ECoG-сигналы имеют лучшее разрешение и большую амплитуду, однако для их получения требуется непосредственное подсоединение электродов к коре головного мозга. Одной из задач при построении систем BCI является предсказание намерений.

Предлагается декодировать исходные сигналы и спрогнозировать траекторию движения верхних конечностей. Исходное пространство имеет избыточно высокую размерность. Линейная зависимость между признаками приводит к мультиколлинеарности. Для устранения мультиколлинеарности предлагается применить методы понижения размерности и отбора признаков.

Признаковое описание многомерного временного ряда существует в пространствах независимых и зависимых переменных. Для учета существующих закономерностей в исходном и выходном пространстве используется скрытое пространство латентных переменных. В скрытом пространстве происходит согласование между образами исходных пространств.

В эксперименте рассматриваются следующие модели: метод частных наименьших квадратов (PLS) [2], отбор признаков с помощью квадратичного программирования (QPFS) [3], метод Белсли (Belsley) и вариации этих методов.

PLS является методом отбора признаков

Описание метода QPFS...

Предлагается система тестирования прогностических моделей с оценкой качества и анализом ошибки. Подобный инструмент может применяться не только в задаче анализа сигналов мозга, но и во многих других задачах, связанных с прогнозированием многомерных временных рядов.

Постановка задачи

Постановка задачи предсказания

* Научный руководитель: Стрижов В. В. Задачу поставил: Стрижов В. В. Консультант: Исаченко Р. В.

Задана выборка $\mathfrak{D} = (\mathbf{X}, \mathbf{Y})$, где $\mathbf{X} \in \mathbb{R}^{m\times n}$ — матрица объектов, $\mathbf{Y} \in \mathbb{R}^{m\times r}$ — матрица ответов. Имеется некоторая модель $\mathbf{f}$ с набором параметров $\boldsymbol{\theta}$ из пространства Θ, предсказывающая $\mathbf{y} \in \mathbb{R}^r$ по $\mathbf{x} \in \mathbb{R}^n$.

Определяется функция ошибки S на выборке $\mathfrak{D}$ и модели $\mathbf{f}$ с параметрами $\boldsymbol{\theta}$. Задачей является поиск наилучших параметров $\boldsymbol{\theta}^*$, то есть таких, при которых функция ошибки минимальна:

$$\boldsymbol{\theta}^* = \arg\min_{\boldsymbol{\theta}\in\Theta} S(\boldsymbol{\theta}|\mathfrak{D}, \mathbf{f}). \tag{1}$$

Однако в случае кореллированных данных $\mathbf{X}$ задача может оказаться нестабильной. Одним из таких случаев является и широко распространенный класс линейных моделей:

$$\mathbf{f}(\mathbf{x}, \boldsymbol{\theta}) = \underset{1\times n}{\mathbf{x}} \cdot \underset{n\times r}{\boldsymbol{\theta}}. \tag{2}$$

За $\mathbf{f}(\mathbf{X}, \boldsymbol{\theta})$ обозначена матрица $[\mathbf{f}(\mathbf{x}_1, \boldsymbol{\theta}), \mathbf{f}(\mathbf{x}_2, \boldsymbol{\theta}), \dots, \mathbf{f}(\mathbf{x}_m, \boldsymbol{\theta})]^\mathsf{T}$. Рассматривается квадратичная функция ошибки:

$$\left\|\mathbf{f}(\mathbf{X}, \boldsymbol{\theta}) - \mathbf{Y}\right\|_2^2 \to \min_{\boldsymbol{\theta}\in\Theta} \tag{3}$$

Если пространство признаков имеет высокую размерность, вероятно, что матрица $\mathbf{X}$ близка к вырожденной, а потому решение проблемы оптимизации (3) будет нестабильным. Поэтому для решения указанной задачи применяются методы отбора признаков, такие как PLS и QPFS.

PLS. Метод частных наименьших квадратов PLS рассматривает в качестве признаков линейные комбинации исходных. Предполагается, что существует скрытое пространство латентных переменных малой размерности l ($l < n, r$). Происходит поиск матрицы $\mathbf{T} \in \mathbb{R}^{m\times l}$, которая наилучшим образом описывает матрицы $\mathbf{X}$ и $\mathbf{Y}$.

$$\underset{m\times n}{\mathbf{X}} = \underset{m\times l}{\mathbf{T}^\mathsf{T}} \cdot \underset{l\times n}{\mathbf{P}^\mathsf{T}} + \underset{m\times n}{\mathbf{E}} \tag{4}$$

$$\underset{m\times r}{\mathbf{Y}} = \underset{m\times l}{\mathbf{U}^\mathsf{T}} \cdot \underset{l\times r}{\mathbf{Q}^\mathsf{T}} + \underset{m\times r}{\mathbf{F}} \tag{5}$$

$\mathbf{T}$ —

QPFS Сформулируем задачу отбора признаков.

Чтобы вычислить матрицу $\mathbf{Q}$ и вектор $\mathbf{b}$

$$\mathrm{Sim} \tag{6}$$

$$\mathrm{Rel} \tag{7}$$

Коэффициент корреляции Пирсона определяется как:

$$\rho_{ij} = \frac{\mathrm{cov}(\mathbf{x}_i, \mathbf{x}_j)}{\sqrt{\mathrm{var}(\mathbf{x}_i) \cdot \mathrm{var}(\mathbf{x}_j)}}$$

Метрики

Пусть имеются истинный прогноз $\mathbf{Y} = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_m)$ и предсказание $\hat{\mathbf{Y}} = (\hat{\mathbf{y}}_1, \hat{\mathbf{y}}_2, \dots, \hat{\mathbf{y}}_m)$; $\mathbf{y}_i$ и $\hat{\mathbf{y}}_i$, $i = 1, 2, \dots, m$ – вектора размерности r.

Среднеквадратичная ошибка (mean squared error):

$$\mathrm{MSE}(\mathbf{Y}, \hat{\mathbf{Y}}) = \frac{1}{m} \sum_{i=1}^{m} \left\| \mathbf{y}_i - \hat{\mathbf{y}}_i \right\|_2^2 \tag{8}$$

Корень среднеквадратичной ошибки (root-mean-squared error):

$$\mathrm{RMSE}(\mathbf{Y}, \hat{\mathbf{Y}}) = \sqrt{\mathrm{MSE}(\mathbf{Y}, \hat{\mathbf{Y}})} \tag{9}$$

Нормированная среднеквадратичная ошибка (scaled mean squared error):

$$\mathrm{sMSE}(\mathbf{Y}, \hat{\mathbf{Y}}) = \frac{\sum_{i=1}^{m} \left\| \mathbf{y}_i - \hat{\mathbf{y}}_i \right\|_2^2}{\sum_{i=1}^{m} \left\| \mathbf{y}_i - \overline{\mathbf{y}} \right\|_2^2},\ \overline{\mathbf{y}} = \frac{1}{m} \sum_{i=1}^{m} y_i \tag{10}$$

MAE, MADE, Коэффициент корреляции. Различные параметры модели из [3].

Литература

[1] *J. del R. Mill?n, F. Renken, J. Mouri?o and W. Gerstner* Brain-actuated interaction // *Artif. Intell.*, 159(2004) 241–259.

[2] *Isachenko R., Vladimirova M., Strijov V.* Dimensionality reduction for time series decoding and forecasting problems // *Machine Learning and Data Analysis.*

[3] *Katrutsa A., Strijov V.* Comprehensive study of feature selection methods to solve multicollinearity problem according to evaluation criteria // *Expert System with Applications*, 76, 1-11.

[4] *Motrenko A., Strijov V.* Multi-way Feature Selection for ECoG-based Brain-Computer Interface // *???*.