Прогнозирование намерений по сигналам ECoG*

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

Работа посвящена построению системы тестирования $L \times K$ прогностических моделей для различных критериев качества. Рассматривается случай коррелированных входных и выходных сигналов электрокортикограммы и фазовых траекторий движения конечностей высокой размерности. 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 ошибку. Вычислительный эксперимент проводится на данных ECoG проекта NeuroTycho.

Ключевые слова: *декодирование временных рядов, PLS, QPFS, электрокортикограмма, траекторий движения конечностей.*

Введение

Работа посвящена исследованию методов моделирования нейросетевого интерфейса (BCI) [5]. Входные данные – сигналы мозга, полученные с помощью электрокортикографии (ECoG) и электроэнцефалографии (EEG). ECoG-сигналы имеют более высокое 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) [4], 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.

---

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

Постановка задачи

Постановка задачи предсказания

Задана выборка $\mathfrak{D} = (\mathbf{X}, \mathbf{Y})$, где $\mathbf{X} \in \mathbb{R}^{m \times n}$ — матрица объектов, $\mathbf{X} = [\chi_1 \dots \chi_n]$, где $\chi_j \in \mathbb{R}^m$ — вектор значений j-го признака на элементах выборки; $\mathbf{Y} \in \mathbb{R}^{m \times r}$ — матрица ответов, $\mathbf{Y} = [y_1 \dots y_r]$. Имеется линейная модель $\mathbf{f}$ с набором параметров $\boldsymbol{\theta}$ из пространства $\mathbb{R}^{n \times r}$, предсказывающая $\mathbf{y} \in \mathbb{R}^r$ по $\mathbf{x} \in \mathbb{R}^n$, следующего вида:

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

Задана 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:

$$\boldsymbol{\theta}^* = \underset{\boldsymbol{\theta} \in \Theta}{\arg\min}\, S(\boldsymbol{\theta} | \mathfrak{D}, \mathbf{f}). \tag{2}$$

В случае кореллированных данных $\mathbf{X}$ решение задачи выбора оптимального вектора параметров $\boldsymbol{\theta}$ нестабильно.

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:

$$\left\| \mathbf{F}(\mathbf{X}, \boldsymbol{\theta}) - \mathbf{Y} \right\|_2^2 \to \min_{\boldsymbol{\theta} \in \mathbb{R}^{n \times r}} \tag{3}$$

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

Постановка задачи отбора признаков

Пусть $\mathcal{J} = \{1, \dots, n\}$ — множество индексов признаков, а $\mathcal{A} \subseteq \mathcal{J}$ — его некоторое подмножество. Тогда линейную модель $\mathbf{f}$ на подмножестве признаков $\mathcal{A}$ можно определить как:

$$\mathbf{f}(\mathbf{x}, \mathcal{A}, \boldsymbol{\theta}) = \underset{1 \times |\mathcal{A}|}{\mathbf{x}_{\mathcal{A}}} \underset{|\mathcal{A}| \times r}{\boldsymbol{\theta}} \tag{4}$$

Функция ошибки S теперь вычисляется на выборке $\mathfrak{D}$, модели $\mathbf{f}$ с параметрами $\boldsymbol{\theta}$ и на $\mathcal{A}$.

Тогда задача предсказания определяется как поиск набора признаков $\mathcal{A}$ и матрицы параметров $\boldsymbol{\theta}^* \in \mathbb{R}^{|\mathcal{A}| \times r}$ такого, что:

$$\boldsymbol{\theta}^* = \underset{\boldsymbol{\theta} \in \mathbb{R}^{|\mathcal{A}| \times r}}{\arg\min}\, S(\theta, \mathcal{A} | \mathfrak{D}, \mathbf{f}) \tag{5}$$

Задача отбора признаков поставлена как:

$$\mathcal{A}^* = \underset{\mathcal{A} \in cJ}{\arg\min}\, Q(\mathcal{A} | \mathbf{X}, \mathbf{y}), \tag{6}$$

где $Q : \mathcal{A} \to \mathbb{R}$ — это некоторый критерий качества,который определяет качество выбранного подмножества признаков $\mathcal{A} \subseteq \mathcal{J}$.

Для решения этой задачи не обязательно требуется оценка оптимального вектора параметров $\boldsymbol{\theta}^*$. Она использует зависимости между векторами $\chi_j, j \in \mathcal{J}$ и целевыми векторами $y_i, i \in \{1, \dots, r\}$.

Пусть $\mathbf{a} \in \mathbb{B}^n = \{0,1\}^n$ – вектор индикаторов принадлежности признаков подмножеству cA – то есть $\mathbf{a}_j = 1$ тогда и только тогда, когда $j \in \mathcal{A}$. Тогда задача 6 может быть записана как:

$$\mathbf{a}^* = \arg\min_{\mathbf{a}\in\mathbb{B}^n} Q(\mathcal{A}|\mathbf{X}, \mathbf{y}), \tag{7}$$

где $Q : \mathbb{B}^n \to \mathbb{R}$ – критерий Q, но определенный на $\mathbb{B}^n$; связь между вектором $\mathbf{a}^*$ и множеством $\mathcal{A}^*$ определяется через:

$$\mathbf{a}^*_j = 1 \Leftrightarrow j \in \mathcal{A}^*, j \in \mathcal{J}. \tag{8}$$

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 $\mathbf{Y}$.

$$\underset{m\times n}{\mathbf{X}} = \underset{m\times l}{\mathbf{T}^{\mathsf{T}}} \cdot \underset{l\times n}{\mathbf{P}^{\mathsf{T}}} + \underset{m\times n}{\mathbf{E}} \tag{9}$$

$$\underset{m\times r}{\mathbf{Y}} = \underset{m\times l}{\mathbf{U}^{\mathsf{T}}} \cdot \underset{l\times r}{\mathbf{Q}^{\mathsf{T}}} + \underset{m\times r}{\mathbf{F}} \tag{10}$$

T –

QPFS

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

$$\mathrm{Sim} \tag{11}$$

$$\mathrm{Rel} \tag{12}$$

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

Среднеквадратичная ошибка (mean squared error):

$$\mathrm{MSE}(\mathbf{Y}, \hat{\mathbf{Y}}) = \frac{1}{m}\sum_{i=1}^{m} \left\|\mathbf{y}_i - \hat{\mathbf{y}}_i\right\|_2^2 \tag{13}$$

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

$$\mathrm{RMSE}(\mathbf{Y}, \hat{\mathbf{Y}}) = \sqrt{\mathrm{MSE}(\mathbf{Y}, \hat{\mathbf{Y}})} \tag{14}$$

Нормированная среднеквадратичная ошибка (scaled mean squared error):

$$\mathrm{sMSE}(\mathbf{Y}, \hat{\mathbf{Y}}) = \frac{\sum_{i=1}^{m} \left\|\mathbf{y}_i - \hat{\mathbf{y}}_i\right\|_2^2}{\sum_{i=1}^{m} \left\|\mathbf{y}_i - \overline{\mathbf{y}}\right\|_2^2},\ \overline{\mathbf{y}} = \frac{1}{m}\sum_{i=1}^{m} y_i \tag{15}$$

MAE, MADE, Коэффициент корреляции. Различные параметры модели из [4].
Сложность модели

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.* Stress test procedure for feature selection algorithms, 2015 // *Expert System with Applications*, 142, 172–183

[4] *Katrutsa A., Strijov V.* Comprehensive study of feature selection methods to solve multicollinearity problem according to evaluation criteria, 2017 // *Expert System with Applications*, 76, 1-11.

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