УДК 621.396:53.08

И.В. БУНЯЕВА, А.А. ЗЕЛЕНСКИЙ

Национальный аэрокосмический университет им. Н.Е. Жуковского «ХАИ», Украина

ОЦЕНКА МАТРИЦЫ РАЗЛОЖЕНИЯ СОВОКУПНОСТИ СИГНАЛОВ НА НЕЗАВИСИМЫЕ КОМПОНЕНТЫ

Методом максимального правдоподобия получен алгоритм формирования оценки матрицы разложения совокупности сигналов на независимые компоненты по наблюдению этой совокупности на фоне нормальных аддитивных помех. Алгоритм получен в виде квадратичной формы, подлежащей минимизации относительно элементов матрицы разложения. Определена потенциальная точность оценок элементов матрицы.

анализ независимых компонент, матрица разложения, шумы наблюдения

Введение

Одной из фундаментальных проблем цифровой обработки сигналов и изображений является поиск адекватных задаче обработки базисов разложения изображений с целью их сжатия, фильтрации и т.п. Чаще всего при этом используются ортогональные преобразования данных (преобразования Фурье, Хаара и т.п.). Каждое из преобразований имеет свои достоинства и недостатки, свою область применения. В ряде случаев очень полезными могут оказаться линейные преобразования, получаемые непосредственно из наблюдаемых данных, наилучшим образом подходящие для обработки конкретной совокупности данных. Разложение на независимые компоненты является одним из таких линейных преобразований. Сущность метода заключается в представлении совокупности наблюдаемых случайных процессов в виде линейной комбинации независимых случайных процессов [1, 2]. Возможно, что наиболее понятной является «проблема состава коктейля», когда имеется большое число ораторов, или источников акустических сигналов, а слушатель слышит смесь этих сигналов. Слушатель имеет только два детектора (уха), и ему трудно выделить (услышать, разобрать) сигнал конкретного, отдельного источника. Очевидно, что проблем будет существенно меньше, если количество детекторов будет равно количеству источников. В электроэнцефалографии, например, наблюдается совокупность колебаний, являющихся взвешенной суммой независимых процессов, генерируемых различными областями мозга. Метод анализа независимых компонент позволяет найти эти независимые компоненты, равно как и соответствующую матрицу разложения. Благодаря этому врач выявляет интересующие его особенности активности мозга.

В течение последних нескольких лет в рамках различных подходов выполнена большая работа по указанной выше проблематике. В большинстве подходов используются методы статистик высокого порядка, минимума совместной информации и максимума энтропии. Кроме метода максимума энтропии, наиболее близкого к обсуждаемой теме, уделено внимание приближению высокого порядка, используемому в популярном методе анализа принципиальных компонент.

Метод *анализа принципиальных компонент (АПК)* используется для разделения совокупности сигналов на т. н. «принципиальные», т.е. некоррелированные, компоненты (преобразование Карунена – Лоэва) [3]. Он основан на том факте, что взаимные кумулянты статистически независимых сигналов равны нулю. В методе АПК (PCA) осуществляется диагонализация матрицы ковариаций, что достигается при некоррелированных сигналах. Поскольку этот метод оперирует только ковариациями, то он учитывает только статистики второго порядка, не используя информацию, содержащуюся в статистиках высокого порядка. Кроме того, при использовании статистик второго порядка для описания сигналов неявно предполагается, что источники являются гауссовыми, но это условие часто не выполняется. Вместе с тем существуют алгебраические методы использования информации, содержащейся в статистиках высокого порядка. Наиболее прямой метод состоит в диагонализации тензора четвертого порядка, составленного из кумулянтов четвертого порядка [3]. Другие методы используют минимизацию взаимной информации, представленной разложением в ряд по моментам высокого порядка или кумулянтам входных сигналов [1, 4].

Термин *анализ независимых компонент (АНК (ICA))* означает поиск статистически независимых источников, соответствующих наблюдаемым данным. Приближение, развитое в [2], использует модель, в которой входная смесь подается на вход нейронной сети с сигмоидальной нелинейностью, преобразующей входной сигнал в выходной. Варьированием весов добиваются максимума совместной энтропии на выходе. Максимизация совместной энтропии на выходе приводит, как ожидается, к минимизации общей информации между выходами, что приводит к их статистической независимости.

Метод максимального правдоподобия использует модель источников с неявными переменными и находит формирующую матрицу, максимизирующую функцию правдоподобия модели для наблюдаемых данных. Показано [5-7], что метод максимального правдоподобия эквивалентен ICA, где форма распределения вероятностей амплитуд источников эквивалентна нелинейностям нейронных сетей из [2]. Аналогичный вывод сделан и в [8-9], где матрица разложения на независимые компоненты ищется максимизацией соответствующей апостериорной плотности вероятности. Вместе с тем существующие алгоритмы разложения на независимые компоненты получены в предположении, что шумы наблюдения отсутствуют. Это существенно сужает область их применения.

В данной статье методом максимума апостериорной плотности вероятности впервые получен алгоритм формирования оценки матрицы разложения совокупности сигналов на независимые компоненты по наблюдению этой совокупности на фоне нормальных аддитивных помех.

1. Формулировка задачи

Колебание, наблюдаемое на выходе i-го датчика на интервале обработки $t \in T$, имеет вид

$$u_i(t) = \sum_{k=1}^{K} A_{ik} s_k(t) + n_i(t), i = \overrightarrow{1, I}, \qquad (1)$$

где A_{ik} – неизвестные элементы матрицы разложения; $s_k(t)$ – неизвестные независимые случайные процессы, энергетический спектр которых равен нулю за пределами полосы частот шириной ΔF. Статистическая независимость случайных процессов $s_k(t)$ означает, что их совместная плотность распределения $p(\mathbf{s}(t))$ равна произведению одномерных плотностей $p_k(s_k(t))$. Положим, что математические ожидания процессов равны нулю, кроме того, они стационарны, эргодичны и имеют одинаковую единичную дисперсию, следовательно,

$$\langle s_k(t) \rangle = 0, \langle s_k(t) s_i(t) \rangle = \delta_{ik},$$

где δ_{ik} – символ Кронекера.

Шумы $n_i(t)$ в различных каналах полагаются статистически независимыми стационарными процессами с равномерным спектром в полосе частот, занятой процессами $s_k(t)$, следовательно,

$$\langle n_i(t) \rangle = 0, \langle n_i(t) n_k(t-\tau) \rangle = \delta_{ik} N_0 \Delta F \frac{\sin \pi \Delta F \tau}{\pi \Delta F \tau}. \qquad (2)$$

Знак $\langle .. \rangle$ означает статистическое усреднение, N_0 –

односторонняя спектральная интенсивность помех.

Задача состоит в том, чтобы разработать алгоритм, обеспечивающий оптимальную (по критерию минимума среднеквадратической ошибки) оценку элементов A_{ik} матрицы разложения и рассчитать теоретически предельную точность оценки.

2. Апостериорная плотность распределения матрицы разложения

При выбранной аппроксимации (ограниченные ширина спектра и интервал наблюдения) процессы (1) с достаточной для практики точностью могут быть представлены конечномерными векторами. Составляющие вектора являются коэффициентами ряда, которым аппроксимируется соответствующий процесс. Так, например, помеха $n_i(t)$ может быть представлена совокупностью $2([\Delta FT]+1)$ независимых нормальных отсчетов $n_i(t/\Delta F)$ и $n_{Ii}(t/\Delta F)$, где $n_{Ii}(t)$ – процесс, сопряженный по Гильберту по отношению к $n_i(t)$, $t=0,1...[\Delta FT]$ [1]. При этом условная функция плотности вероятности совокупности колебаний $\{u_i(t)\}$, $i=\overline{1,I}$, представляемых матрицей отсчетов $\mathbf{u}^t$, при условии, что $\mathbf{A}=\|A_{ik}\|$ и неизвестные процессы $s_k(t)$ (или матрица $\mathbf{s}^t$ их отсчетов) имеют данное фиксированное значение (функция правдоподобия параметров $\mathbf{A}$ и $\mathbf{s}^t$), выражается формулой [1]

$$p(\mathbf{u}^t/\mathbf{A},\mathbf{s}^t)=$$

$$=k\exp\left\{-\frac{1}{N_0}\sum_{i=1}^{I}\sum_{t=0}^{m}\left[u_i^t-\sum_{k=1}^{K}A_{ik}s_k^t\right]^2\right\}, \qquad (3)$$

где $k=(2\pi\Delta FN_0)^{-(m+1)I}$, $m=[\Delta FT]$ – число отсчетов по времени.

Апостериорная совместная плотность вероятности матриц $\mathbf{A}$ и $\mathbf{s}^t$ определяется формулой Байеса

$$p(\mathbf{A},\mathbf{s}^t/\mathbf{u}^t)=p(\mathbf{u}^t/\mathbf{A},\mathbf{s}^t)p(\mathbf{A})p(\mathbf{s}^t). \qquad (4)$$

Поскольку о композиционной матрице $\mathbf{A}$ априори ничего не известно, то соответствующее априорное распределение примем равномерным для всех возможных композиционных матриц:

$$p(\mathbf{A})=\text{const},$$

в результате получим

$$p(\mathbf{A},\mathbf{s}^t/\mathbf{u}^t)=p(\mathbf{u}^t/\mathbf{A},\mathbf{s}^t)p(\mathbf{s}^t). \qquad (5)$$

В принципе, максимизацией уравнения (5) по $\mathbf{A}$ и $\mathbf{s}$ могут быть найдены оценки как композиционной матрицы, так и неизвестных сигналов. Однако размерность этой задачи очень велика и ее решение требует значительных вычислительных ресурсов.

Обычно типовые алгоритмы поиска композиционной матрицы и вектора неизвестных сигналов (алгоритмы анализа независимых компонент) получают в отсутствие тепловых шумов. Максимизация (5) выполняется при этом один раз для одного временного сечения наблюдаемых процессов, что снижает размерность задачи, но уменьшает устойчивость этого решения к шумам наблюдения.

Другим способом размерность может быть снижена и в условиях наличия шумов наблюдения, если положить на первом этапе $\mathbf{s}^t$ несущественным параметром для того, чтобы усреднить апостериорную плотность (5) по этому параметру. В результате получаем апостериорное распределение матрицы $\mathbf{A}$

$$p(\mathbf{A}/\mathbf{u}^t)=\int p(\mathbf{u}^t/\mathbf{A},\mathbf{s}^t)p(\mathbf{s}^t)d\mathbf{s}^t. \qquad (6)$$

Далее в случае необходимости можно подставить оценку $\mathbf{A}$ (6) в (5) и найти оценку $\mathbf{s}$. Воспользуемся статистической независимостью сигналов s_j^t ($p(s^t)$ может быть заменено произведением плотностей вероятностей независимых источников $p_j(s_j^t)$). Для простоты положим, что независимыми являются и все временные сечения этих сигналов. Используя также соотношение (3), преобразуем (6) к следующему виду:

$$p(\mathbf{A}/\mathbf{u}^t)=\int..\int k\exp\left\{-\frac{1}{N_0}\sum_{i=1}^{I}\sum_{t=0}^{m}\left[u_i^t-\sum_{k=1}^{K}A_{ik}s_k^t\right]^2\right\}\times$$

$$\times\prod_t\prod_j p_j(s_j^t)ds_j^t\,. \qquad (7)$$

Многомерный интеграл (7) может быть представлен произведением одномерных интегралов:

$$p(\mathbf{A}/\mathbf{u}^t)=\prod_t\int\prod_i k\exp\left\{-\frac{1}{N_0}\left[u_i^t-\sum_{k=1}^{K}A_{ik}s_k^t\right]^2\right\}\times$$
$$\times\prod_j p_j(s_j^t)ds_j^t\,. \qquad (8)$$

Содержащиеся в (8) интегралы могут быть вычислены методом асимптотического интегрирования (методом перевала) [11], при этом

$$p(\mathbf{A}/\mathbf{u}^t)=\prod_t(\det A)^{-1}\prod_j p_j({A^{-1}}_{jk}u_k^t)\,. \qquad (9)$$

Для определения максимума логарифмируем последнее соотношение. Имеем

$$\ln p(\mathbf{A}/\mathbf{u}^t)=-m\ln\det A+$$
$$+\sum_t\sum_j\log p_j({A^{-1}}_{jk}u_k^t)+C\,, \qquad (10)$$

где C – логарифм коэффициента пропорциональности, неявно присутствующего в (10).

Для упрощения обозначим $\mathbf{W}=\mathbf{A}^{-1}$, при этом

$$\ln p(\mathbf{A}/\mathbf{u}^t)=$$
$$=m\ln\det W+\sum_t\sum_j\ln p_j(W_{jk}u_k^t)+C. \qquad (11)$$

3. Алгоритм формирования оценки матрицы разложения

Наилучшая оценка композиционной матрицы $\mathbf{A}$ (или однозначно связанной с ней матрицы $\mathbf{W}$) определяется как координата максимума логарифма апостериорной плотности $p(\mathbf{A}/\mathbf{u}^t)$. Для нахождения соответствующего алгоритма формирования оценки матрицы разложения приравняем к нулю производную логарифма апостериорной плотности $p(\mathbf{A}/\mathbf{u}^t)$ по элементам матрицы $\mathbf{W}$. Интересующая нас производная равна

$$\frac{\partial}{\partial W_{ji}}\ln p(\mathbf{A}/\mathbf{u}^t)=$$
$$=\frac{\partial}{\partial W_{ji}}\left[m\ln\det W+\sum_t\sum_j\ln p_j(W_{jk}u_k^t)+C\right]=$$
$$=mA_{ij}+\frac{\partial}{\partial W_{ji}}\left[\sum_t\sum_j\ln p_j(W_{jk}u_k^t)\right]=$$
$$=mA_{ij}+\sum_t\frac{\partial x_j}{\partial W_{ji}}\frac{\partial}{\partial x_j}\sum_j\ln p_j(x_j^t)=$$
$$=mA_{ij}+\sum_t\frac{\partial x_j^t}{\partial W_{ji}}\frac{\partial\ln p_j(x_j^t)}{\partial x_j^t}=$$
$$=mA_{ij}+\sum_t\left[u_j^t\left(\frac{p_j'(W_{jk}u_k^t)}{p_j(W_{jk}u_k^t)}\right)_j\right]=$$
$$=mA_{ij}+\sum_t\left[u_j^t\left(l(W_{jk}u_k^t)\right)_j\right]. \qquad (12)$$

Перепишем (12) в матричной форме:

$$\frac{\partial}{\partial W}\ln \mathrm{p}(\mathbf{A}/\mathbf{u}^t)=mW^{-T}+\sum_t\left(l(W_{jk}u_k^t)\right)_j\mathbf{u}^{tT}. \qquad (13)$$

В (12) и (13) символом $l(W_{jk}u_k^t)$ обозначена производная логарифма априорной плотности распределения независимых компонент, т. е. $l(x)=d/dx[\ln p(x)]$. Причем, это, в общем случае, нелинейное преобразование $l(x)$, примененное к вектору-столбцу x, дает новый вектор-столбец, элементы которого получены нелинейным преобразованием соответствующих элементов исходного вектора. Следующее из (13) матричное нелинейное уравнение правдоподобия

$$W^{-T}+\frac{1}{T}\sum_t l(W\,\mathbf{u}^t)\mathbf{u}^{tT}=0 \qquad (14)$$

в явном виде может быть решено только для одного частного случая нормального априорного распределения, когда $l(x)=-x$ (единичная дисперсия априорного распределения). При этом

$$W^{-T}-\frac{1}{T}\sum_t W\,\mathbf{u}^t\mathbf{u}^{tT}=$$
$$=W^{-T}-W\frac{1}{T}\sum_t\mathbf{u}^t\mathbf{u}^{tT}=W^{-T}+WR_U=0. \qquad (15)$$

Умножая левую и правую части уравнения (15) на W^{1}, находим

$$W = R_U^{-1/2} \tag{16}$$

и, соответственно, учитывая, что $W = A^{-1}$, имеем

$$A = R_U^{1/2}. \tag{17}$$

В (15) - (17) символом R_U обозначена матрица коэффициентов корреляции элементов вектора входных данных u^t:

$$R_U = \frac{1}{T}\sum_t \mathbf{u}^t \mathbf{u}^{tT}. \tag{18}$$

Решение (16), (17) может быть использовано как нулевое приближение при решении уравнения (14) методом последовательных приближений. При этом итерационная процедура строится следующим образом:

1) определяется с помощью (16) нулевое приближение W_0;

2) первое приближение находится подстановкой W_0 в (14):

$$W_i^{-T} = \frac{1}{T}\sum_t l(W_{i-1}\mathbf{u}^t)\mathbf{u}^{tT}; \tag{19}$$

3) найденное решение обращается, транспонируется, и далее повторяется шаг (2).

Итерационный процесс останавливается, если норма разности решений, полученных на смежных шагах, оказывается меньше некоторой достаточно малой величины ε, т.е.

$$\|W_{i-1} - W_i\| \le \varepsilon.$$

В условиях надежных измерений (малый уровень шумов), в окрестности истинных значений элементов матрицы разложения $\alpha_{ik} = \alpha_{ik}^0$ выражение (11) хорошо аппроксимируется первыми членами разложения в ряд Тейлора. Точность аппроксимации тем выше, чем больше отношение сигнал/шум:

$$L(\mathbf{A}) = L(\mathbf{A}^0) + L'(\mathbf{A}^0)(\mathbf{A} - \mathbf{A}^0) + \\ + \frac{1}{2}L''(\mathbf{A}^0)(\mathbf{A} - \mathbf{A}^0)^2 + \ldots$$

При этом формула для оценки $\bar{\mathbf{A}}$ значений элементов матрицы разложения упрощается:

$$\bar{\mathbf{A}} = \mathbf{A}^0 - L'(\mathbf{A}^0) / L''(\mathbf{A}^0).$$

Легко показать, что в условиях надежных измерений дисперсия оценки $\bar{\mathbf{A}}$ определяется соотношением

$$\sigma_{\mathbf{A}}^2 = \langle |L'(\mathbf{A}^0)|^2 \rangle / |L''(\mathbf{A}^0)|^2.$$

Таким образом, при условии, что априорное распределение $\mathbf{s}(t)$ подчиняется нормальному закону, матрица $\mathbf{R_u}$ раскладывает исследуемый процесс в ряды Карунена – Лоэва, что подтверждает правильность используемого подхода.

Заключение

Прежде всего отметим, что полученные результаты являются предварительными, требуют конкретизации и экспериментальной проверки. Вместе с тем, в отличие от [1-7] полученные выше оценки строятся не на основании одного временного сечения совокупности наблюдаемых процессов, а с использованием достаточно мощных статистик R_U, что само по себе должно гарантировать их хорошую устойчивость и качество.

Литература

1. Comon P. Independent component analysis, a new concept? // Signal Processing, 36. –1994. – P. 287–314.

2. Pearlmutter B.A, Parra L.C. A context-sensitive generalization of ICA // International Conference on Neural Information Processing – Hong Kong. – 1996. – P. 507 – 515.

3. Cardoso J.-F. A tetradic decomposition of 4th-order tensors: application to the source separation problem, in M. Moonen and B. de Moor (eds.), Algorithms, Architectures and Applications // Volume III of SVD and Signal Processing, Elsevier. – Amsterdam. – 1995. – P. 375 – 382.

4. Bell A.J., Sejnowski T.J. An information-maximization approach to blind separation and blind deconvolution // Neural Computation, 7. – 1995. – P. 1129 – 1159.

5. Cardoso J.-F. Infomax and maximum likelihood for source separation // IEEE Letters on Signal Processing, 4(4). – 1997. – P. 112 – 114.

6. Knuth K.H. Difficulties Applying Recent Blind Source Separation Techniques to EEG and MEG, in Maximum Entropy and Bayesian Methods // Boise, Idaho, USA. – 1997. – P. 209 – 222.

7. Analysis and Visualization of Single-Trial Event-Related Potentials /Jung T.P, Makeig S., Westerfield M., Townsend J., Courchesne E. and Sejnowski T.J. // Human Brain Mapping 14. – 2001. – P. 166 – 185.

8. Hyvarinen A. Survey on independent component analysis // Neural Computing Surveys, 2. – 1999. – P. 94 – 128.

9. Фалькович С.Е., Пономарев В.И., Шкварко Ю.В. Оптимальный прием пространственно-временных сигналов в радиоканалах с рассеянием. – М.: Радио и связь, 1989. – 296 с.

10. Де Брейн. Асимптотические методы в анализе. – М.: Иностранная литература, 1961. – 354 с.

11. Mutihac R., Van Hulle M.M. Neural network implementations of Independent Component Analysis // Proceedings of the IEEE Workshop. – 2002. – P. 505 – 514.

12. Classification and ICA using Maximum Likelihood Hebbian Learning / R. Corchado, J. Koetsier, D. Macdonald, C. Fyte // Proceedings of the IEEE Workshop. – 2002. – P. 327 – 336.

13. Bell J., Sejnowski T.J. An information-maximization approach to blind separation and blind deconvolution // Neural Computation, 7. – 1995. – P. 1129 – 1159.

14. Pham T., Garrat P., Jutten C. Separation of a mixture of independent sources through a maximum likelihood approach // Proc. EUSIPCO. – 1992. – P. 771 – 774.

15. A unifying information-theoretic framework for Independent Component Analysis / W. Lee, M. Girolami, A.J. Bell, T.J. Sejnowski // Computers and Mathematics with applications. – 2000. – Vol. 39. – P. 1 – 21.

16. Monsher J.C., Leahy R.M., Lewis P.S. EEG and MEG: Forward Solutions for Inverse Methods // IEEE Transaction On Biomedical Engineering. – 1999. – Vol. 46, № 3. – P. 245 – 259.

Поступила в редакцию 05.05.04

Рецензент: д-р техн. наук, ст. науч. сотр. В.В. Лукин, Национальный аэрокосмический университет им. Н.Е. Жуковского "ХАИ", г. Харьков