Д. Р. Дорошин, **Д. А. Леднов** (Москва, Компания «Стел–КС»). **Выявление регрессионных зависимостей между компонентами векторов признаков речи.**

В системах обработки речи в качестве векторов признаков, как правило, используются MFCC (Mel-frequency cepstral coefficients) и PLP (The Perceptual Linear Predictive) [1]. Обычно предполагается, что компоненты векторов признаков независимы. Однако специфика методов предобработки такова, что это предположение не является оправданным. В докладе рассмотрена регрессионная модель этой зависимости и предложена методика нахождения участков зависимости между компонентами.

Пусть $o(t) = [o_0(t), \ldots, o_N(t)]^T$ $(t = 1, \ldots, M)$ — последовательность векторов признаков, где $N+1$ — размерность вектора, t — время. Если предположить регрессионную зависимость на всем интервале измерений, то уравнение регрессии можно представить в следующем виде:

$$o_0(t) = a_1 o_1(t) + \cdots + a_N o_N(t) + r(t), \qquad t = 1, \ldots, M, \tag{1}$$

где $a_1, \ldots, a_N$ — коэффициенты регрессии, а $r(t)$ — ошибка аппроксимации.

Допустим, что весь интервал измерений можно поделить на подынтервалы, на каждом из которых своя регрессионная зависимость. В данном случае уравнение (1) можно переписать в следующем виде:

$$o_0(t) = a_1^{s(t)} o_1(t) + \cdots + a_N^{s(t)} o_N(t) + r(t), \quad t \in \{1, \ldots, M\}. \tag{2}$$

Здесь $s(t)$ — номер соответствующего набора коэффициентов. Задача состоит в определении всех наборов $\{a_j^s\}$, $j \in \{1, \ldots, N\}$, $s \in \{1, \ldots, K\}$, где K — количество наборов.

Для решения задачи предположим, что поведение каждого коэффициента описывается гауссовским случайным процессом с независимыми приращениями. Вторым этапом является аппроксимация полученных коэффициентов регрессии кусочно-постоянными функциями, что позволит определить все возможные наборы коэффициентов (2) и моменты переходов между ними.

Учитывая данные предположения, задачу можно переписать в виде

$$\begin{aligned} &a_i(t+1) = a_i(t) + q_i(t), \qquad i = 1, \ldots, N, \\ &o_0(t) = a_1 o_1(t) + \cdots + a_N o_N(t) + r(t), \qquad t = 1, \ldots, M. \end{aligned}$$

Здесь $q(t) = \{q_j(t)\}$ и $r(t)$ — белые гауссовские шумы с нулевыми математическими ожиданиями и ковариационными матрицами Q и R.

Принятая модель переменных коэффициентов $a_1(t), \ldots, a_N(t)$ позволяет поставить задачу оценивания в рамках калмановского подхода. Алгоритм оценивания основан на минимизации дисперсии ошибки оценки [2]. Результатом алгоритма оценивания являются оценки коэффициентов регрессии $\widetilde{a}_j(t)$ как функций времени.

Второй этап задачи состоит в определении всех наборов $\{a_j^s\}$ (2) по полученным оценкам $\widetilde{a}_1(t), \ldots, \widetilde{a}_N(t)$. Эта задача может быть решена при помощи известного EM-алгоритма (Expectation-Maximization algorithm) [1].

СПИСОК ЛИТЕРАТУРЫ

1. *Аграновский А. В.*, *Леднов Д. А.* Теоретические аспекты алгоритмов обработки и классификации речевых сигналов. М.: Радио и связь, 2004.
2. *Dan Simon.* Optimal State Estimation. John Wiley & Sons, Inc., 2006.