Выбор функций потерь в задачах неотрицательного матричного разложения

Рябенко Евгений Алексеевич

Диссертация на соискание учёной степени
кандидата физико-математических наук
05.13.18 — математическое моделирование,
численные методы и комплексы программ

Научный руководитель — д.ф.-м.н. К.В. Воронцов

ВЦ РАН, 30 октября 2014 г.

Неотрицательное матричное разложение

Дано: матрица $P \in \mathbb{R}_+^{m \times n}$, число $r < \min(m, n)$.

Найти: матрицы $A \in \mathbb{R}_+^{m \times r}$, $X \in \mathbb{R}_+^{r \times n}$ такие, что

$$P \approx AX \equiv Q.$$

Оптимизационная задача:

$$(A^*, X^*) = \underset{A \geqslant 0,\, X \geqslant 0}{\operatorname{argmin}} D(P, AX).$$

Дивергенция — неотрицательная сепарабельная функция потерь с единственным нулём:

$$D(P, Q) = \sum_{i=1}^{m} \sum_{j=1}^{n} d(p_{ij}, q_{ij}),$$

$$d(p, q) \geqslant 0,$$

$$d(p, q) = 0 \Leftrightarrow p = q.$$

Примеры прикладных задач

1. Рекомендательные системы

$$R_{iu} = \sum_t p_{it} q_{tu}$$

дано: R_{iu} — рейтинги товаров i, поставленные пользователем u;
найти: p_{it} — профиль интересов товара i;
q_{tu} — профиль интересов пользователя u.

2. Тематическое моделирование текстовых коллекций

$$f_{wd} = \sum_t \phi_{wt} \theta_{td}$$

дано: f_{wd} — частоты слов w в документах d;
найти: ϕ_{wt} — распределения слов w в темах t;
θ_{td} — распределения тем t в документах d.

3. Анализ данных ДНК-микрочипов

$$I_{pk} = \sum_g a_{pg} c_{gk}$$

дано: I_{pk} — интенсивность флуоресценции p-й пробы на k-м чипе;
найти: a_{pg} — коэффициент сродства p-й пробы g-му гену;
c_{gk} — уровень экспрессии g-го гена на k-м чипе.

Разновидности дивергенций

$$D(P,Q) = \sum_{i=1}^{m}\sum_{j=1}^{n} d(p_{ij}, q_{ij}), \quad d(p,q) \geqslant 0, \quad d(p,q) = 0 \Leftrightarrow p = q.$$

Дивергенция	$d(p,q)$
норма l_1	$d_1(p,q) = \lvert p-q \rvert$
квадрат нормы Фробениуса	$d_F(p,q) = (p-q)^2$
дивергенция Кульбака-Лейблера	$d_{KL}(p,q) = p\ln\frac{p}{q} - p + q$
дивергенция Итакура-Саито	$d_{IS}(p,q) = \ln\frac{q}{p} + \frac{p}{q} - 1$
расстояние Хеллингера	$d_H(p,q) = \left(\sqrt{p} - \sqrt{q}\right)^2$
χ^2 Пирсона	$d_P(p,q) = \frac{(p-q)^2}{q}$
χ^2 Неймана	$d_N(p,q) = d_P(q,p) = \frac{(p-q)^2}{p}$

Минимизация некоторых дивергенций эквивалентна максимизации правдоподобия в известных параметрических моделях:

Дивергенция	Модель шума	$p(P)$
Фробениуса	аддитивная гауссовская	$\prod_{ij} N\left(p_{ij}, \sigma^2\right)$
Кульбака-Лейблера	пуассоновская	$\prod_{ij} P(p_{ij})$
Итакура-Саито	мультипликативная гамма	$\prod_{ij} G(p_{ij}, \alpha/q_{ij})$

Различные дивергенции оптимальны для разных моделей шума.

Семейство АБ-дивергенций

АБ-дивергенция [Cichocki, 2011]:

$$d_{AB}^{(\alpha,\beta)}(p,q) = \begin{cases} \frac{1}{\alpha\beta}\left(\frac{\alpha}{\alpha+\beta}p^{\alpha+\beta} + \frac{\beta}{\alpha+\beta}q^{\alpha+\beta} - p^{\alpha}q^{\beta}\right), & \alpha,\beta,\alpha+\beta \neq 0, \\ \frac{1}{\alpha^2}\left(p^{\alpha}\ln\frac{p^{\alpha}}{q^{\alpha}} - p^{\alpha} + q^{\alpha}\right), & \alpha \neq 0,\ \beta = 0, \\ \frac{1}{\alpha^2}\left(\ln\frac{q^{\alpha}}{p^{\alpha}} + \left(\frac{q^{\alpha}}{p^{\alpha}}\right)^{-1} - 1\right), & \alpha = -\beta \neq 0, \\ \frac{1}{\beta^2}\left(q^{\beta}\ln\frac{q^{\beta}}{p^{\beta}} - q^{\beta} + p^{\beta}\right), & \alpha = 0,\ \beta \neq 0, \\ \frac{1}{2}\left(\ln p - \ln q\right)^2, & \alpha = \beta = 0. \end{cases}$$

α регулирует разреженность модели,
β определяет соотношение между эффективностью и устойчивостью получаемых оценок.

Идея работы: когда модель шума неизвестна, задачу выбора оптимальной функции потерь можно свести к выбору параметров в семействе АБ-дивергенций.

Задача оптимизации гиперпараметров α и β

Пусть существует семейство плотностей $p(P, \alpha, \beta)$ вида

$$p(P, \alpha, \beta) = \frac{1}{Z(\alpha, \beta)} p_0(P, \alpha, \beta),$$

$$p_0(P, \alpha, \beta) = e^{-D_{AB}^{(\alpha,\beta)}(P, Q(P))} = \prod_{i,j} e^{-d_{AB}^{(\alpha,\beta)}(p_{ij}, q_{ij})}, \tag{1}$$

$$Z(\alpha, \beta) = \int_X p_0(X, \alpha, \beta)\, dX.$$

Тогда оценка максимального правдоподобия для α и β имеет вид

$$(\alpha^*, \beta^*) = \operatorname*{argmax}_{\alpha,\beta} \sum_{i,j} \ln p(p_{ij}, \alpha, \beta).$$

Проблема: нормировочный множитель $Z(\alpha, \beta)$ не выражается аналитически или даже не существует (интеграл расходится).

Метод согласования вклада (score matching)

Поскольку нормировочный множитель $Z(\alpha, \beta)$ неизвестен, вместо метода максимизации правдоподобия можно использовать метод согласования вклада [Hyvärinen, 2006, 2007].

Пусть $p_T(x)$ — истинная плотность распределения данных,
$p(x, \theta)$ — модельное семейство плотностей.

ОМП (оценка максимума правдоподобия):

$$\theta^* = \operatorname*{argmin}_\theta \int_x p_T(x) \ln \frac{p_T(x)}{p(x, \theta)} dx.$$

ОСВ (оценка согласования вклада):

$$\theta^* = \operatorname*{argmin}_\theta \int_x p_T(x) \left\| \nabla_x \ln \frac{p_T(x)}{p(x, \theta)} \right\|^2 dx.$$

$\nabla_x \ln p(x, \theta) = \nabla_x \ln p_0(x, \theta) \Longrightarrow$
согласование вклада можно использовать, не зная $Z(\alpha, \beta)$.
Более того, не обязательно даже, чтобы он существовал [Hyvärinen, 2008].

Метод оптимизации гиперпараметров α и β

Теорема

В модели (1) *OCB принимает следующий вид:*

$$(\alpha^*, \beta^*) = \underset{\alpha,\beta}{\operatorname{argmin}} J(P, \alpha, \beta),$$

$$J(P, \alpha, \beta) = \begin{cases} \frac{1}{\beta} \sum\limits_{i,j} p_{ij}^{\alpha} \left(\frac{1}{2\beta} p_{ij}^{\alpha} \left(p_{ij}^{\beta} - q_{ij}^{\beta} \right)^2 - p_{ij}^{\beta} (\alpha + \beta + 1) + q_{ij}^{\beta} (\alpha + 1) \right), \\ \qquad \beta \neq 0, \\ \sum\limits_{i,j} p_{ij}^{\alpha} \left(\ln \frac{q_{ij}}{p_{ij}} \left(\frac{p_{ij}^{\alpha}}{2} \ln \frac{q_{ij}}{p_{ij}} + \alpha + 1 \right) - 1 \right), \\ \qquad \beta = 0. \end{cases}$$

Оптимальные значения α и β предлагается находить, численно решая приведённую задачу минимизации.

Задача неотрицательного матричного разложения с фиксированной функцией потерь

Оптимизационная задача в общем виде:

$$(A^*, X^*) = \underset{A \geqslant 0,\ X \geqslant 0}{\operatorname{argmin}} D\left(P, AX\right).$$

$D\left(P, AX\right)$ не выпукла по совокупности аргументов, поэтому используются блочно-покоординатные методы минимизации:

Вход: матрица P, ранг разложения r;
Выход: матрицы-множители A и X;

1 инициализация $A^0 \geqslant 0,\ X^0 \geqslant 0$;
2 **для всех** итераций $t = 1, 2, \ldots$
3 $\quad X^t = f\left(P, A^{t-1}, X^{t-1}\right)$;
4 $\quad \left(A^t\right)^T = f\left(P^T, \left(X^t\right)^T, \left(A^{t-1}\right)^T\right)$.

Чаще всего используются мультипликативные алгоритмы, позволяющие естественным образом сохранять неотрицательность элементов матриц.

Задача неотрицательного матричного разложения с фиксированной АБ-дивергенцией

Оптимизационная задача для АБ-дивергенции:

$$(A^*, X^*) = \underset{A \geqslant 0,\, X \geqslant 0}{\operatorname{argmin}} D_{AB}^{(\alpha,\beta)}(P, AX). \tag{2}$$

Мультипликативный алгоритм для АБ-дивергенции [Cichocki, 2011]:

$$X \leftarrow X \otimes \left(\left(A^T\left(P^{[\alpha]} \otimes Q^{[\beta-1]}\right)\right) \oslash \left(A^T Q^{[\alpha+\beta-1]}\right)\right)^{[\omega(\alpha,\beta)]},$$

$$A \leftarrow A \otimes \left(\left(\left(P^{[\alpha]} \otimes Q^{[\beta-1]}\right) X^T\right) \oslash \left(Q^{[\alpha+\beta-1]} X^T\right)\right)^{[\omega(\alpha,\beta)]},$$

$$\omega(\alpha,\beta) = \begin{cases} \frac{1}{1-\beta}, & \frac{\beta}{\alpha} < \frac{1}{\alpha} - 1, \\ \frac{1}{\alpha}, & \frac{\beta}{\alpha} \in \left[\frac{1}{\alpha} - 1, \frac{1}{\alpha}\right], \\ \frac{1}{\alpha+\beta-1}, & \frac{\beta}{\alpha} > \frac{1}{\alpha}. \end{cases}$$

$\otimes$ — поэлементное умножение матриц, $\oslash$ — поэлементное деление, $Z^{[z]}$ — поэлементное возведение матрицы Z в степень z.

Правая часть — глобальный минимум квадратичной функции, мажорирующей $D_{AB}^{(\alpha,\beta)}$ на текущей итерации; следовательно, в ходе обновлений функция потерь монотонно невозрастает.

Сходимость мультипликативного алгоритма

Поскольку задача не является выпуклой, лучшее, что можно гарантировать — сходимость к стационарной точке, задаваемой условиями Каруша-Куна-Таккера:

$$\begin{cases} X^* \geqslant 0, \\ \nabla_X D_{AB}^{(\alpha,\beta)}\left(P, A^*X^*\right) \geqslant 0, \\ X^* \otimes \nabla_X D_{AB}^{(\alpha,\beta)}\left(P, A^*X^*\right) = 0, \\ A^* \geqslant 0, \\ \nabla_A D_{AB}^{(\alpha,\beta)}\left(P, A^*X^*\right) \geqslant 0, \\ A^* \otimes \nabla_A D_{AB}^{(\alpha,\beta)}\left(P, A^*X^*\right) = 0. \end{cases}$$

Проблема: обновления мультипликативного алгоритма могут останавливаться в нестационарных точках вблизи нулей: если $x_{kj} = 0$, то он останется равным нулю, даже если $\left[\nabla_X D_{AB}^{(\alpha,\beta)}\right]_{kj} < 0$.

ε-модификация мультипликативного алгоритма

Отделим A и X от нуля небольшой положительной константой ε:

$$\begin{aligned} X &\leftarrow \max\left(\varepsilon, X \otimes \left(\left(A^T\left(P^{[\alpha]} \otimes Q^{[\beta-1]}\right)\right) \oslash \left(A^T Q^{[\alpha+\beta-1]}\right)\right)^{[\omega(\alpha,\beta)]}\right), \\ A &\leftarrow \max\left(\varepsilon, A \otimes \left(\left(\left(P^{[\alpha]} \otimes Q^{[\beta-1]}\right) X^T\right) \oslash \left(Q^{[\alpha+\beta-1]} X^T\right)\right)^{[\omega(\alpha,\beta)]}\right). \end{aligned} \tag{3}$$

Теорема

При любом $\varepsilon > 0$ функция $D_{AB}^{(\alpha,\beta)}(P, AX)$ монотонно невозрастает при обновлениях (3) *для любого начального приближения $A^0 \geqslant \varepsilon,\ X^0 \geqslant \varepsilon$.*

Теорема

Алгоритм с обновлениями (3) *для любого начального приближения $A^0 \geqslant \varepsilon,\ X^0 \geqslant \varepsilon$ сходится к стационарной точке отделённой от нуля задачи*

$$(A_\varepsilon^*, X_\varepsilon^*) = \underset{A \geqslant \varepsilon,\, X \geqslant \varepsilon}{\operatorname{argmin}} D_{AB}^{(\alpha,\beta)}(P, AX).$$

Метод ε-прореживания матриц A^*_ε и X^*_ε

Получив решение отделённой от нуля задачи, проредим его, обнулив элементы, равные ε:

$$X = X^*_\varepsilon \otimes [X^*_\varepsilon > \varepsilon],$$
$$A = A^*_\varepsilon \otimes [A^*_\varepsilon > \varepsilon].$$

Теорема

Для матриц (A, X), полученных из $(A^*_\varepsilon, X^*_\varepsilon)$ ε-прореживанием, верно следующее: $\forall\, i, k, j$

$$\begin{cases} \left[\begin{array}{ll} a_{ik} = 0, & [\nabla_A D_{AB}]_{ik} \geqslant -\mathcal{O}(\varepsilon), \\ a_{ik} > 0, & \left|[\nabla_A D_{AB}]_{ik}\right| \leqslant \mathcal{O}(\varepsilon), \end{array}\right. \\ \left[\begin{array}{ll} x_{kj} = 0, & [\nabla_X D_{AB}]_{kj} \geqslant -\mathcal{O}(\varepsilon), \\ x_{kj} > 0, & \left|[\nabla_X D_{AB}]_{kj}\right| \leqslant \mathcal{O}(\varepsilon), \end{array}\right. \end{cases}$$

то есть, в точке (A, X) условия стационарности исходной задачи (2) выполняются с точностью до $\mathcal{O}(\varepsilon)$.

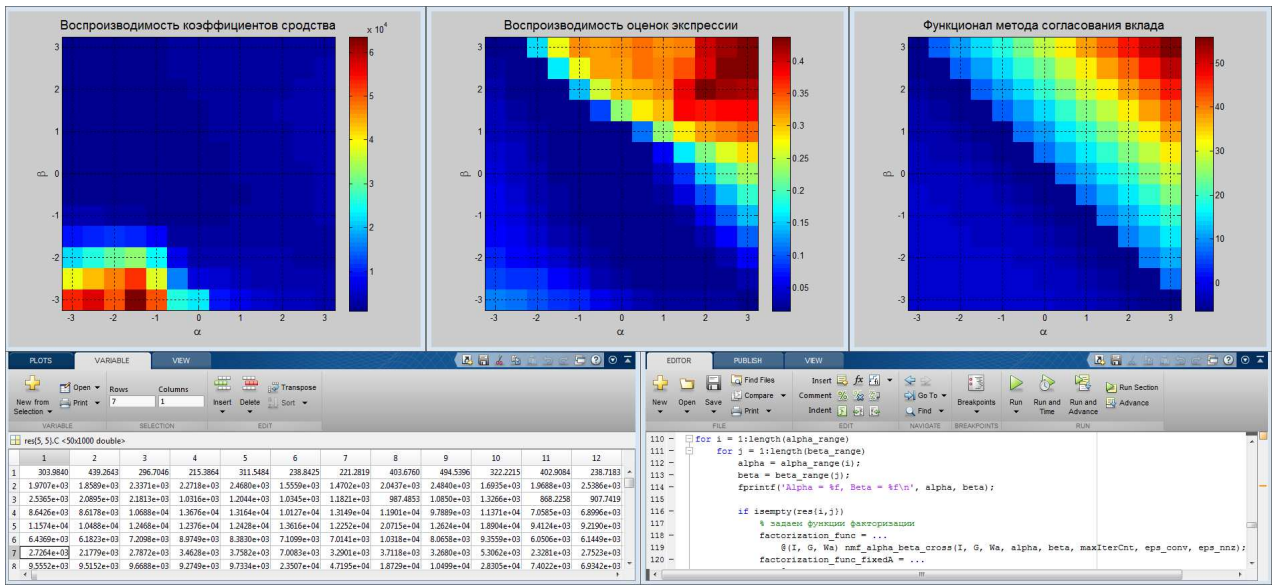
Результаты

- Для задачи неотрицательного матричного разложения предложен способ выбора оптимальной функции потерь в семействе АБ-дивергенций, основанный на методе согласования вклада.
- Предложен ε-модифицированный мультипликативный алгоритм неотрицательного матричного разложения с АБ-дивергенцией. Показано, что:
  - в ходе его применения функция потерь монотонно невозрастает;
  - алгоритм глобально сходится к стационарной точке оптимизационной задачи, отделённой от нуля;
  - модификация решения, полученная ε-прореживанием, даёт точку, условия стационарности исходной задачи в которой выполняются с точностью до $\mathcal{O}(\varepsilon)$.

Практическая часть работы

Разработанные методы неотрицательного матричного разложения были применены к задаче анализа данных ДНК-микрочипов.

- Предложен ряд моделей, основанных на неотрицательном матричном разложении и учитывающих особенности данных, игнорируемые стандартными методами анализа.
- Создан программный комплекс, представляющий собой библиотеку модулей и средств визуализации для адаптивного неотрицательного матричного разложения, настройки предложенных моделей и анализа микрочиповых экспериментов.

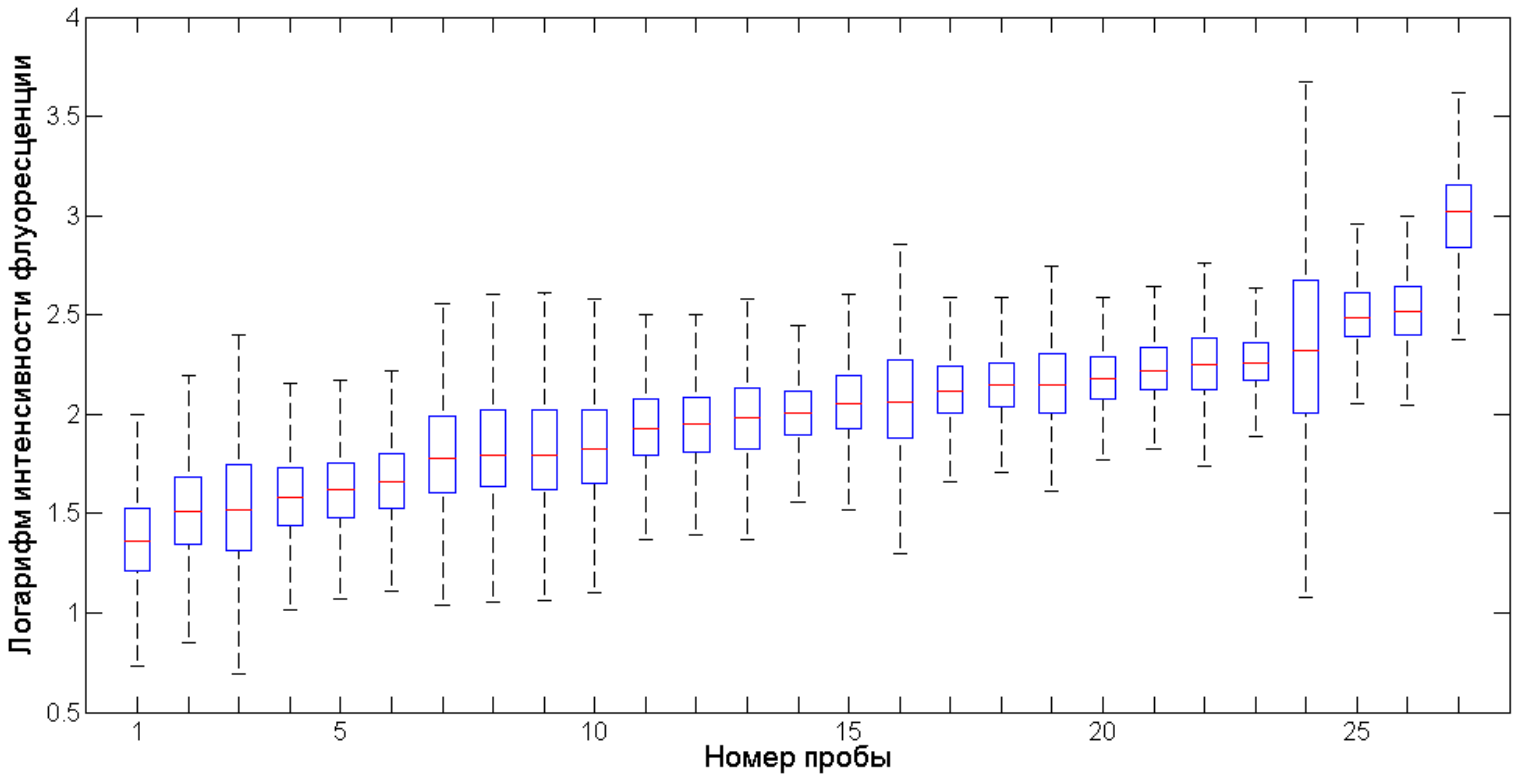


Особенности задачи анализа данных ДНК-микрочипов

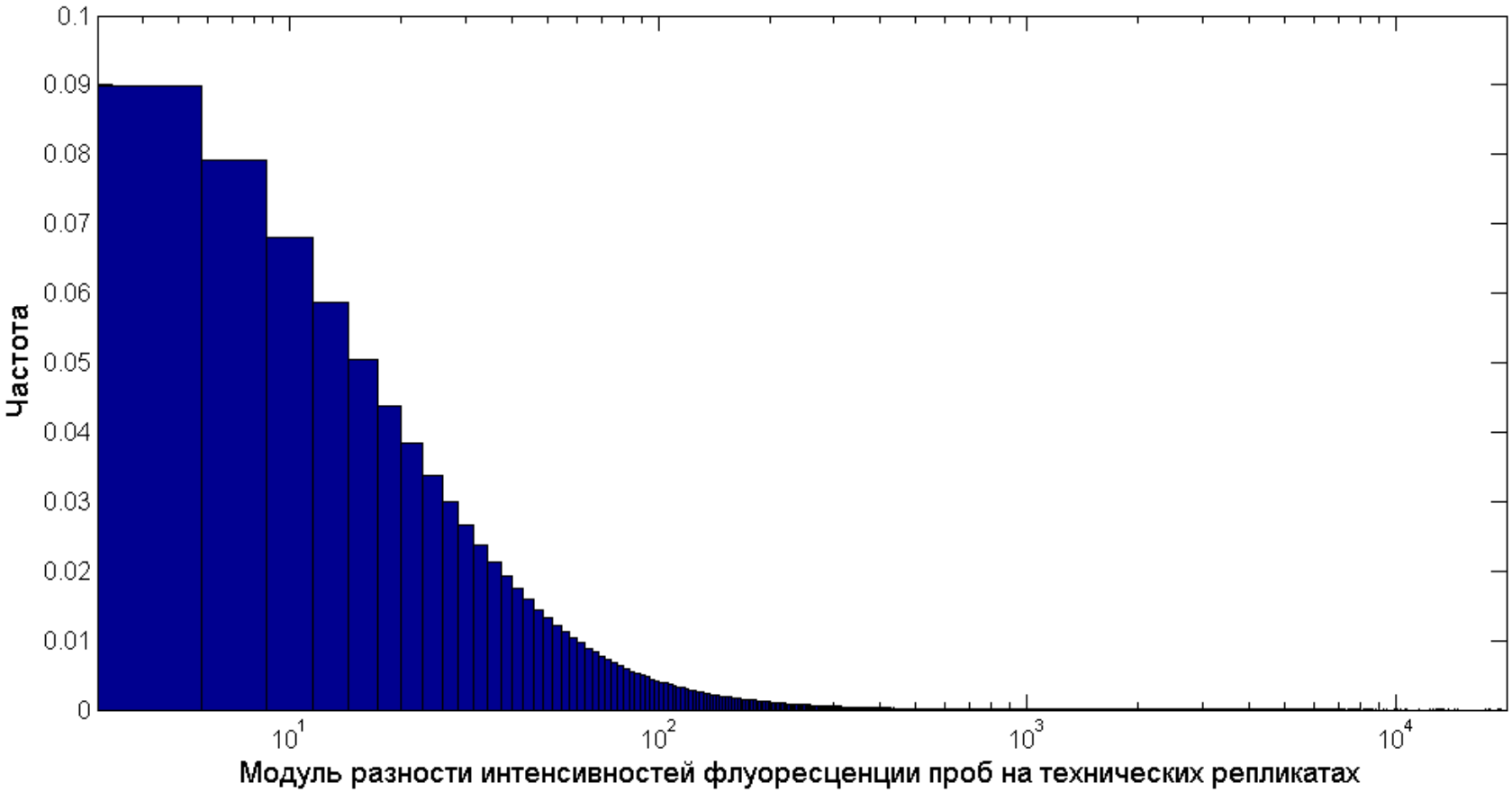
ДНК-микрочип — устройство, позволяющее оценивать экспрессию десятков тысяч генов одновременно.

Экспрессия каждого гена оценивается с помощью десятков флуоресцирующих сенсоров — **проб**.

Проблема: систематические различия между показаниями проб одного и того же гена, вызванные их физическими свойствами:



Особенности задачи анализа данных ДНК-микрочипов



Распределение шума имеет тяжёлые хвосты и не может быть оценено непосредственно — не существует экспериментов с известным сигналом.

Модель, учитывающая степени сродства проб с геном

Известные данные:
I_{pk} — интенсивность флуоресценции пробы p на микрочипе k;
$g(p)$ — номер гена, для которого проба p **специфична**
(определён конструкцией микрочипа).

Неизвестные параметры:
c_{gk} — уровень экспрессии гена g на микрочипе k;
a_p — коэффициент **сродства** (affinity) пробы p гену $g(p)$.

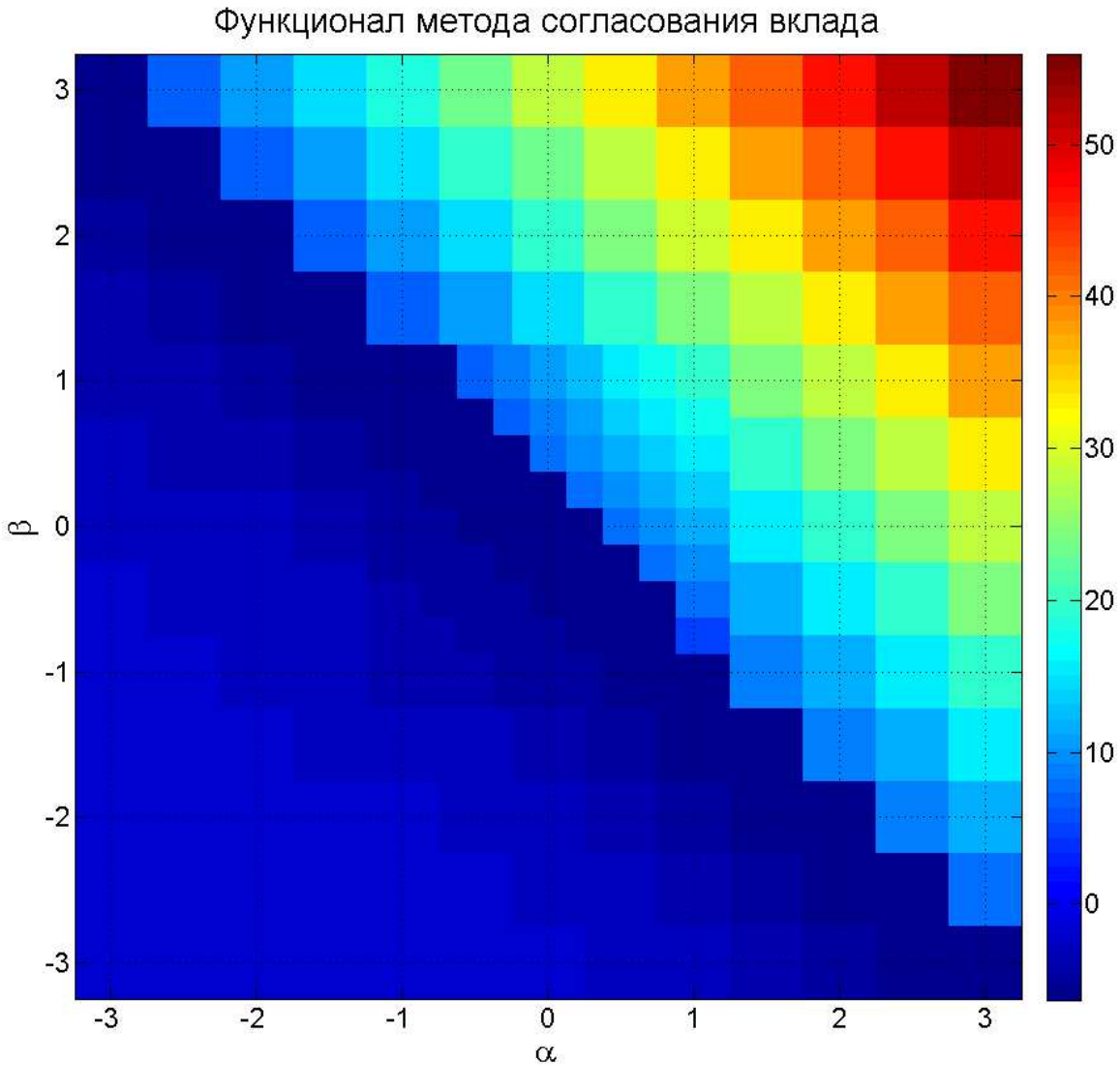
$$I_{pk} \approx \hat{I}_{pk} = a_p c_{g(p)k}.$$

Задача распадается на G независимых подзадач неотрицательного матричного разложения ранга 1.

В стандартных методах анализа микрочипов используется именно такая модель, но коэффициенты сродства в ней не фиксированы,
а определяются каждый раз по анализируемой выборке.

База данных GEO содержит данные тысяч микрочиповых экспериментов.
Идея: использовать эту информацию для настройки моделей.

Результаты оценки качества модели как функции от α и β



Функционал достигает минимума при $\alpha = -0.5,\ \beta = 0.75$
(традиционная логнормальная модель шума соответствует $\alpha = \beta = 0$).

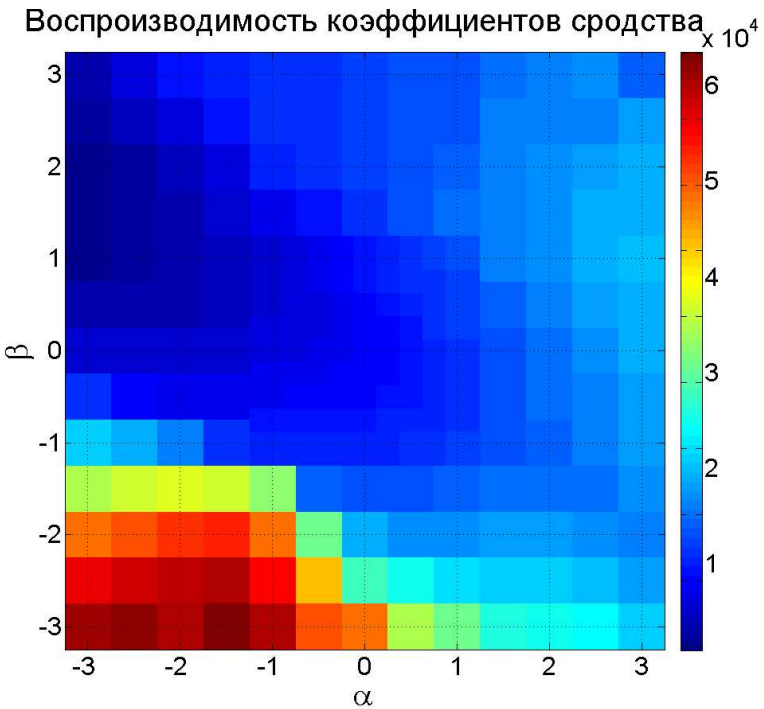
Результаты оценки качества модели как функции от α и β

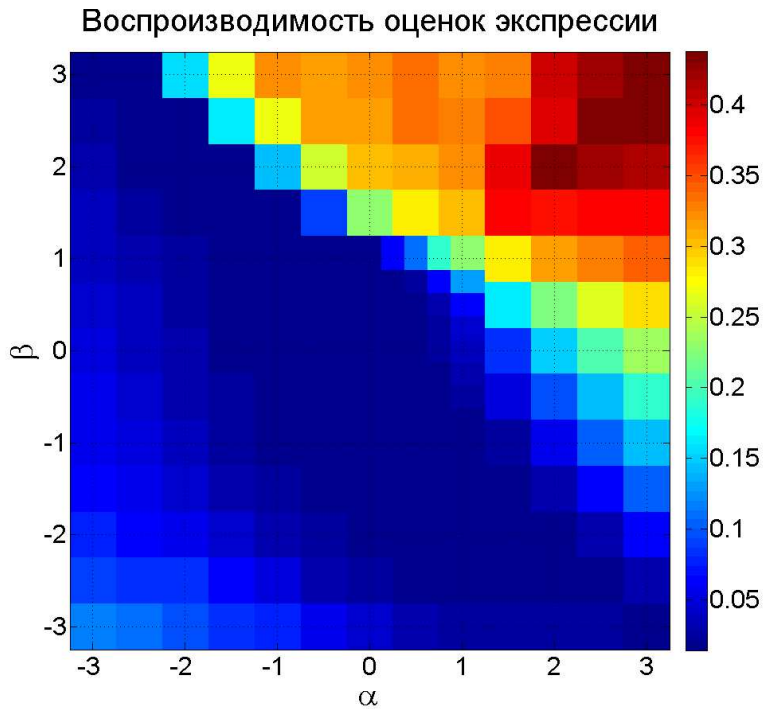
Воспроизводимость коэффициентов сродства по двум подмножествам чипов:

$$rep_a = \frac{1}{G}\sum_{g=1}^{G}\frac{1}{P(g)}\sum_{p\in P(g)}\frac{\left|a_{pg}^1 - a_{pg}^2\right|}{a_{pg}^1 + a_{pg}^2}.$$

Воспроизводимость оценок экспрессии:

$$rep_c = \frac{1}{G}\sum_{g=1}^{G}\frac{1}{K}\sum_{k=1}^{K}\frac{\left|c_{gk}^1 - c_{gk}^2\right|}{c_{gk}^1 + c_{gk}^2}.$$

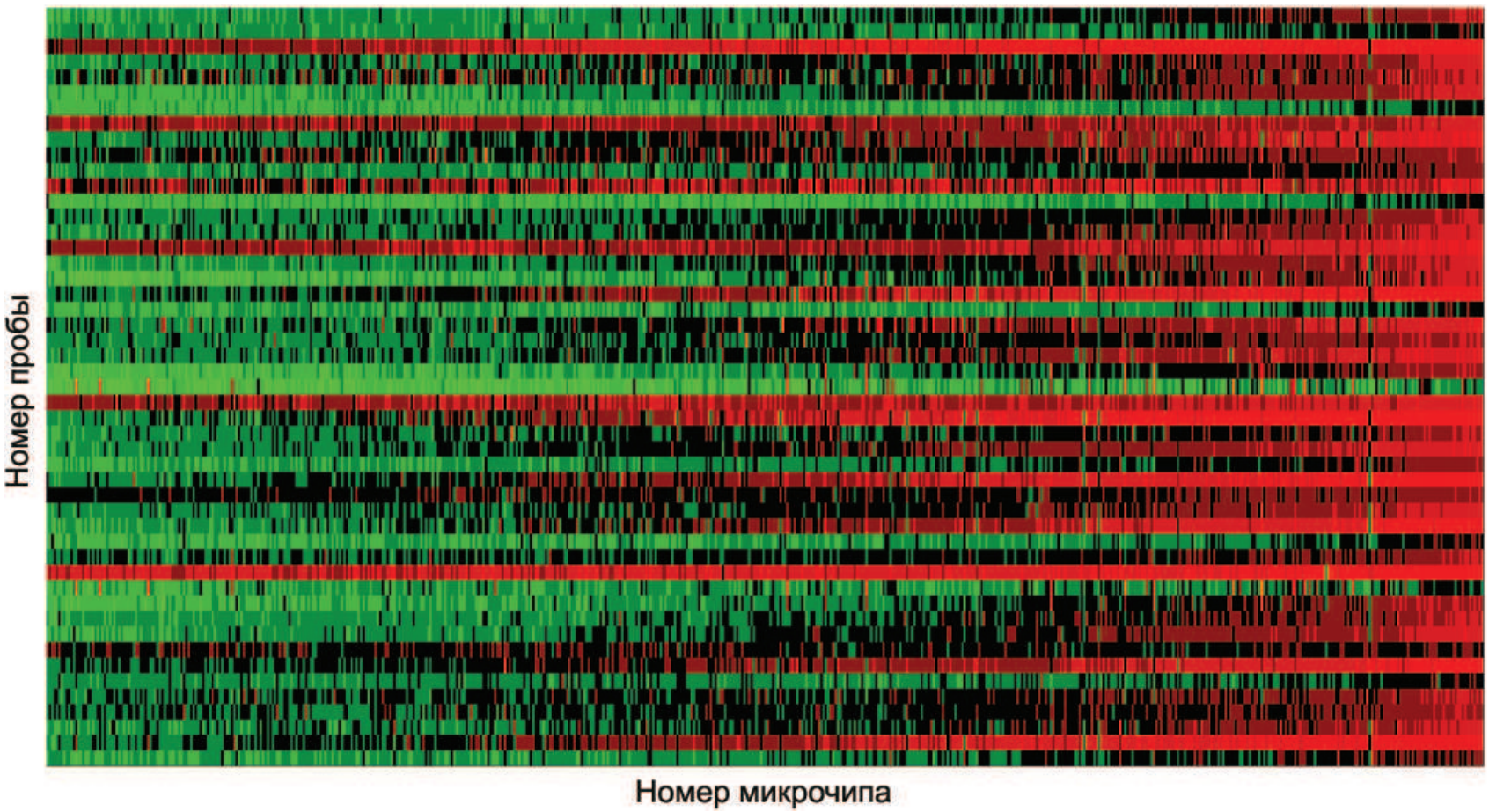




Эффект альтернативного сплайсинга

Альтернативный сплайсинг — экспрессия **части** гена.

Проблема: пробы к отсутствующим частям не оценивают экспрессию:



Модель, учитывающая эффект альтернативного сплайсинга

Идея: после настройки модели рассчитаем относительную ошибку, пропорциональную концентрациям и обратно пропорциональную интенсивностям:

$$e_{pk} = \frac{\hat{I}_{pk} - I_{pk}}{I_{pk}} \cdot c_{g(p)k}.$$

Пусть $e_{0.95}$ — 95% выборочный квантиль e_{pk}; создадим матрицу бинарных весов $W \in \{0,1\}^{P \times K}$ с элементами $w_{pk} = [e_{pk} < e_{0.95}]$.

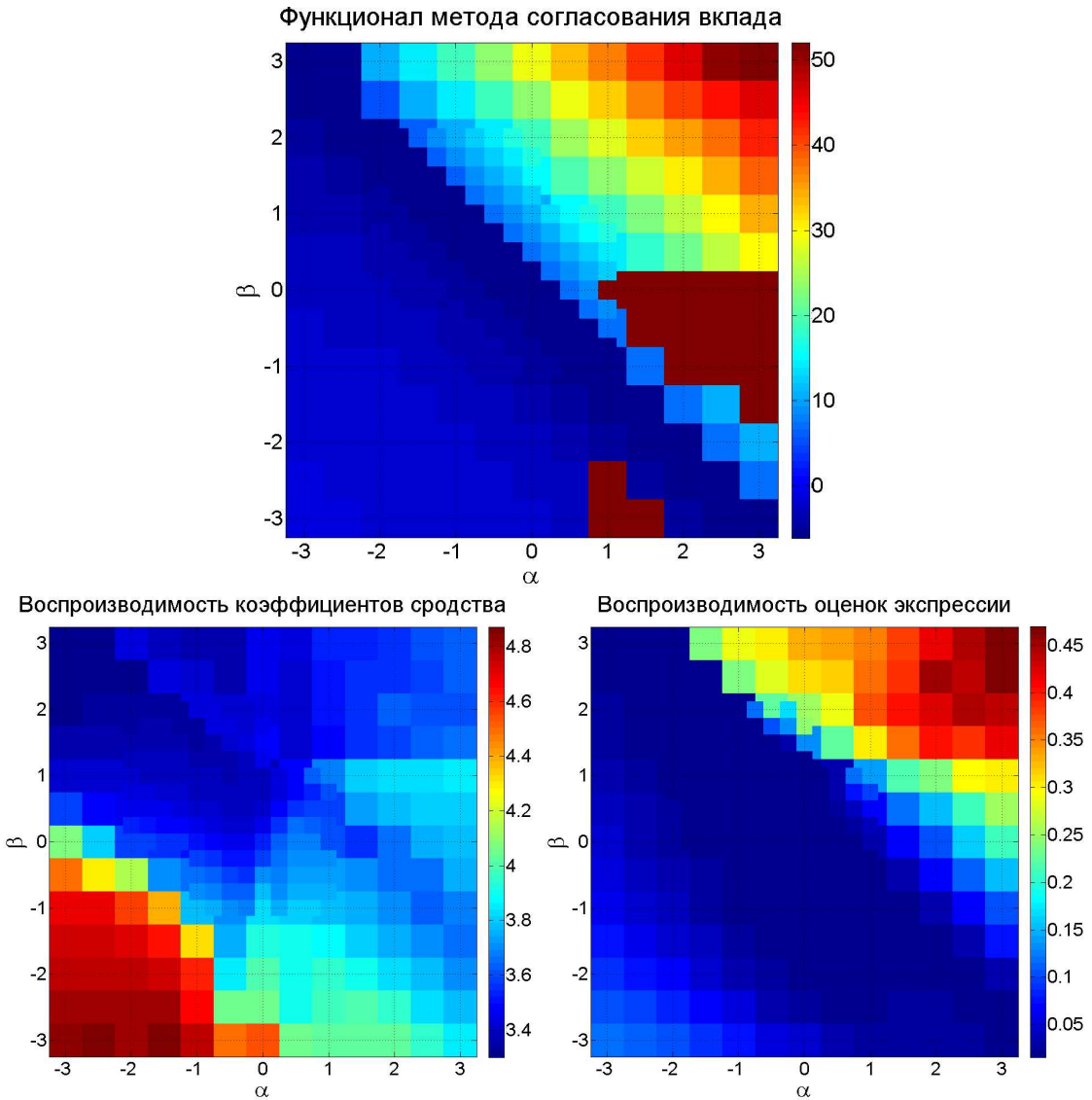
Веса легко встраиваются в обновления мультипликативного алгоритма:

$$X \leftarrow \max\left(\varepsilon, X \otimes \left(\left(A^T\left(P^{[\alpha]} \otimes Q^{[\beta-1]} \otimes W\right)\right) \oslash \left(A^T\left(Q^{[\alpha+\beta-1]} \otimes W\right)\right)\right)^{[\omega(\alpha,\beta)]}\right),$$

$$A \leftarrow \max\left(\varepsilon, A \otimes \left(\left(\left(P^{[\alpha]} \otimes Q^{[\beta-1]} \otimes W\right)X^T\right) \oslash \left(\left(Q^{[\alpha+\beta-1]} \otimes W\right)X^T\right)\right)^{[\omega(\alpha,\beta)]}\right).$$

Будем повторять настройку модели и обновление весов несколько раз.

Результаты оценки качества модели как функции от α и β



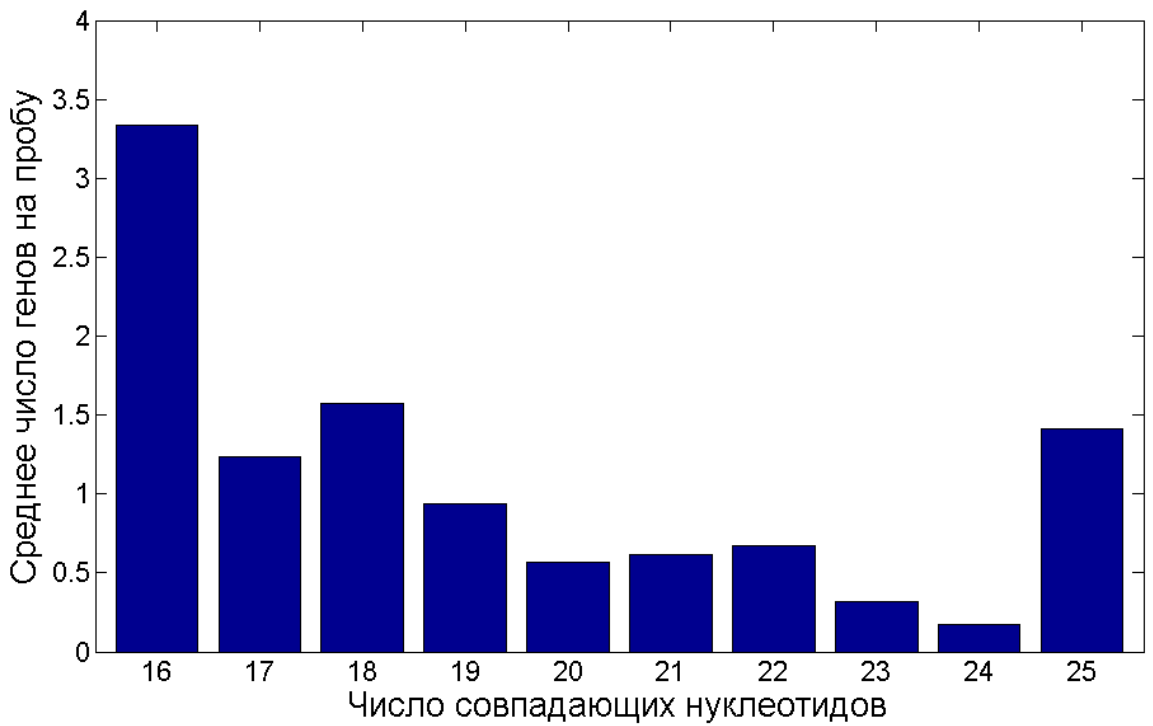
Функционал метода согласования вклада достигает минимума при $\alpha = -0.75, \quad \beta = 0.75.$

Эффект кросс-гибридизации

Свечение пробы может быть вызвано генами, для которых она не специфична.

Проба 432:309		C	T	G	C	C	A	C	A	T	T	G	C	T	G	A	G	G	C	T	C	A	G	A	G	C	
Ген GRIA1	...	G	A	C	G	G	T	G	T	A	A	C	G	A	C	T	C	C	G	A	G	T	C	T	C	G	...
Ген GRIA3	...	G	A	C	G	G	T	G	T	A	A	C	G	A	G	T	C	C	G	A	G	T	C	T	C	G	...
Ген SNRPN	...	G	A	C	G	G	T	G	T	G	A	C	G	A	C	T	C	C	T	A	G	T	C	C	A	C	...
Ген DNAJC22	...	G	A	C	G	G	T	G	T	A	T	C	G	A	C	T	C	C	A	C	C	C	A	G	A	T	...

Распределение среднего числа комплементарных генов:



Модель, учитывающая эффект кросс-гибридизации

Рассмотрим факторизованную модель ранга G:

$$I_{pk} \approx \hat{I}_{pk} = \sum_{g=1}^{G} a_{pg} c_{gk}.$$

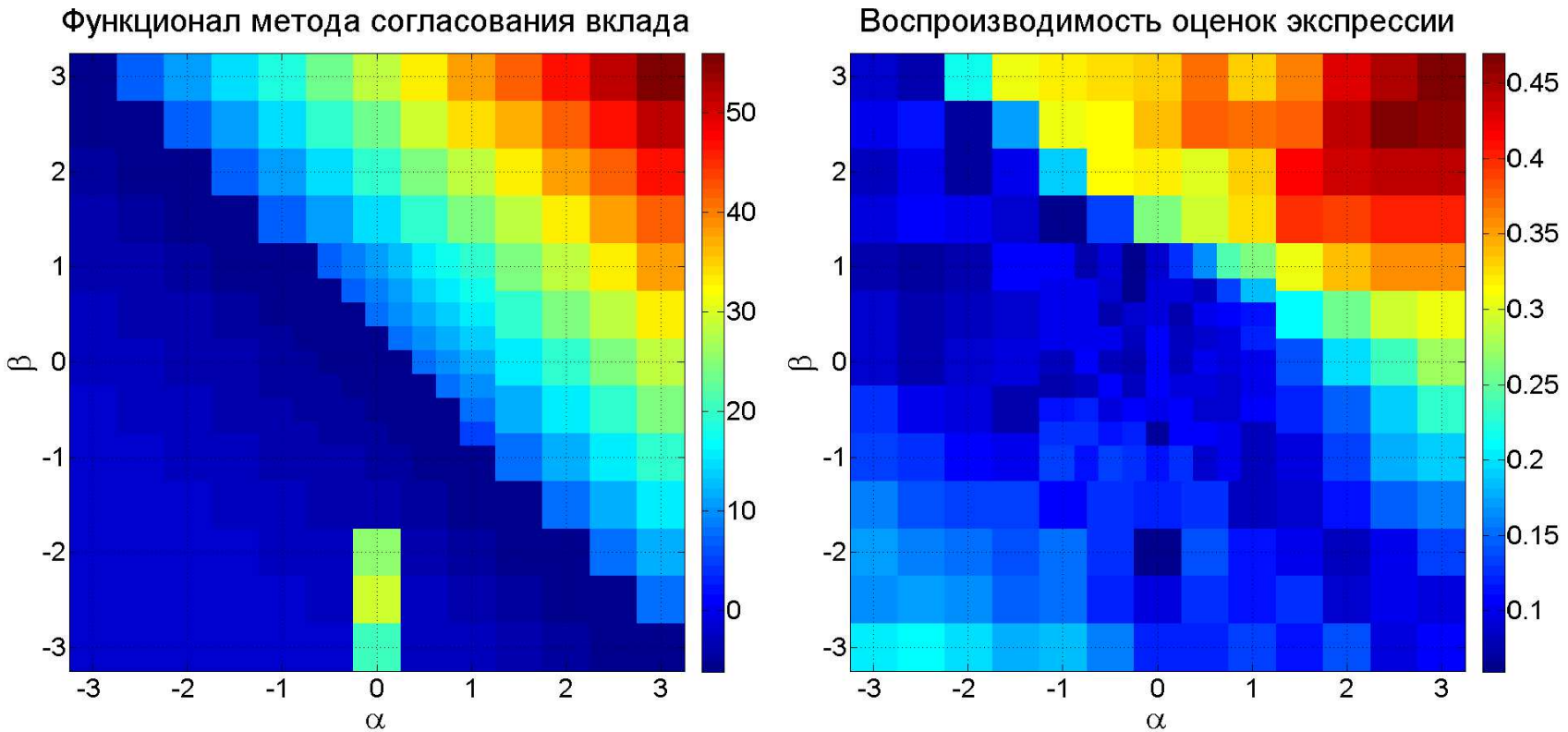
Используем информацию о сходстве последовательностей проб и генов: положим $a_{pg} = 0$, если n_{pg} — число совпадающих нуклеотидов в пробе p и гене g — меньше 20.

Сформируем матрицу бинарных весов $W \in \{0,1\}^{P \times G}$ с элементами $w_{pg} = [n_{pg} \geqslant 20]$.

Веса встраиваются в обновления матрицы A:

$$A \leftarrow W \otimes \max\left(\varepsilon, A \otimes \left(\left(\left(P^{[\alpha]} \otimes Q^{[\beta-1]}\right) X^T\right) \oslash \left(Q^{[\alpha+\beta-1]} X^T\right)\right)^{[\omega(\alpha,\beta)]}\right).$$

Результаты оценки качества модели как функции от α и β



Функционал метода согласования вклада достигает минимума при $\alpha = -0.5, \quad \beta = 0.75.$

Дополнительные критерии качества

Используем данные эксперимента [Affymetrix, 2007], в котором смеси двух образцов в известных пропорциях наносились на несколько чипов каждая:

№ смеси j	1	2	3	4	5	6	7	8	9
Число чипов n_j	3	3	3	3	9	3	3	3	3
Доля РНК мозга p_j	0	0.05	0.1	0.25	0.5	0.75	0.9	0.95	1
Доля РНК сердца $1 - p_j$	1	0.95	0.9	0.75	0.5	0.25	0.1	0.05	0

Вариабельность оценок экспрессии между повторами экспериментов (средний квадрат чистых ошибок):

$$var_{mix} = \frac{1}{27G} \sum_{g=1}^{G} \sum_{j=1}^{9} \sum_{u=1}^{n_j} \left(c_{gu}^{j} - \bar{c}_{g}^{j} \right)^2 .$$

Степень нелинейности оценок экспрессии (средний квадрат ошибок, обусловленных неадекватностью модели):

$$lin_{mix} = \frac{1}{9G} \sum_{g=1}^{G} \sum_{j=1}^{9} n_j \left(\hat{c}_{g}^{j} - \bar{c}_{g}^{j} \right)^2 .$$

Значения дополнительных критериев качества на полученных моделях

Метод	Учитываемые эффекты			var_{mix}	lin_{mix}
	постоянство коэффициентов сродства	альтернативный сплайсинг	кросс-гибридизация		
RMA	-	-	-	386.6	189.4
1	+	-	-	160.4	114.1
2	+	+	-	154.4	119.4
3	+	-	+	168.0	117.4

RMA — наиболее популярный метод оценки экспрессии.

Результаты

- Предложенный метод выбора оптимальной функции потерь в множестве АБ-дивергенций применён в задаче анализа экспрессии генов с помощью ДНК-микрочипов; получены оценки неизвестного распределения шума.
- Предложены следующие модели и методы их настройки:
  - модель, учитывающая постоянство коэффициентов сродства;
  - модель, учитывающая эффект альтернативного сплайсинга;
  - модель, учитывающая эффект кросс-гибридизации.
- Проведены эксперименты, показывающие, что каждая из настроенных моделей позволяет уменьшить вариабельность оценок экспрессии между повторами экспериментов на 56-60% и степень нелинейности оценок экспрессии — на 37-40%.

Публикации

1. Рябенко, Е. А. (2014). Мультипликативный метод неотрицательного матричного разложения с АБ-дивергенцией и его сходимость. Машинное обучение и анализ данных, 1(7), 800–816.
2. **Крайнова, Н. А., Хаустова, Н. А., Макеева, Д. С., Федотов, Н. Н., Гудим, Е. А., Рябенко, Е. А., Шкурников, М. Ю., Галатенко, В. В., Сахаров, Д. А., Мальцева, Д. В. (2013). Оценка потенциальных референсных генов для нормализации данных ПЦР-РВ в экспериментах с клетками линии HeLa. Биотехнология, 1, 42–50.**
3. **Рябенко, Е. А. (2012). Настройка нелинейной модели данных экспериментов с экспрессионными ДНК-микрочипами. Математическая биология и биоинформатика, 7(2), 554–566.**
4. **Sakharov, D. A., Maltseva, D. V, Riabenko, E. A., Shkurnikov, M. U., Northoff, H., Tonevitsky, A. G., Grigoriev, A. I. (2012). Passing the anaerobic threshold is associated with substantial changes in the gene expression profile in white blood cells. European journal of applied physiology, 112(3), 963–972.**
5. Riabenko, E., Kogadeeva, M., Gavrilyuk, K., Sokolov, E., Shanin, I., Tonevitsky, A. G. (2012). Comparing Affymetrix Human Gene 1.0 ST preprocessing methods on tissue mixture data. 6th International Conference on Bioinformatics and Biomedical Engineering (iCBBE) (pp. 631–634). Shanghai, China.
6. **Мальцева, Д. В., Рябенко, Е. А., Сизова, С. В., Яшин, Д. В., Хаустова, С. А., Шкурников, М. Ю. (2012). Влияние физической нагрузки на экспрессию генов HSPBP1, PGLYRP1 и HSPA1A в лейкоцитах человека. Бюллетень экспериментальной биологии и медицины, 153(6), 846–850.**
7. **Riabenko, E. A., Tonevitsky, E. A., Tonevitsky, A. G., Grigoriev, A. I. (2011). Structural Peculiarities of Human Genes Which Expression Increases in Response to Stress. American Journal of Biomedical Sciences, 3(2), 90–94.**
8. Рябенко, Е. А., Когадеева, М. С. (2011). Нижняя граница числа комплементарных нуклеотидов при моделировании кросс-гибридизации. ММРО-15, г. Петрозаводск. (с. 540–542). МАКС Пресс.
9. Когадеева, М. С., Рябенко, Е. А. (2011). Математическая модель данных микрочипов ДНК, учитывающая эффекты кросс-гибридизации и насыщения. ММРО-15, г. Петрозаводск. (с. 536–539). МАКС Пресс.

Результаты, выносимые на защиту

- Метод адаптивного выбора функционала потерь в задаче неотрицательного матричного разложения.
- Метод получения неотрицательного матричного разложения с АБ-дивергенцией в качестве функции потерь, доказательство его глобальной сходимости к точке, сколь угодно близкой к стационарной.
- Модели данных экспериментов с ДНК-микрочипами, учитывающие коэффициенты сродства, эффекты альтернативного сплайсинга и кросс-гибридизации, настроенные с помощью метода адаптивного выбора функционала потерь.
- Комплекс программ, позволяющий получить оценки экспрессии генов на основе предложенных моделей.