Khnum: быстрая open-source программа для расчета метаболических потоков с использованием ^{13}C-углерода

Стешин Семен Сергеевич

МГУ ВМК, кафедра математической кибернетики, 2020

Научный руководитель:
к.ф.м.н., доцент
Шуплецов М. С.

Анализ Метаболических Потоков

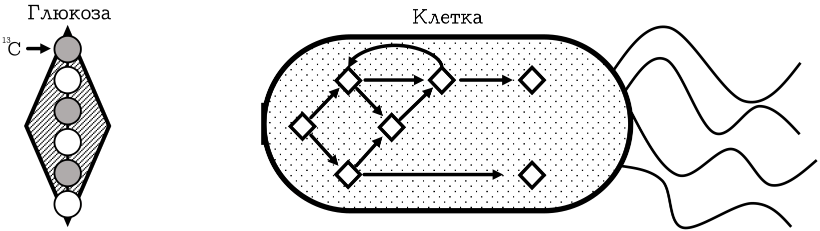
Метаболический поток — внутриклеточная химическая реакция

^{13}C-Metabolic Flux Analysis (*^{13}C-MFA*) — метод измерения метаболических потоков

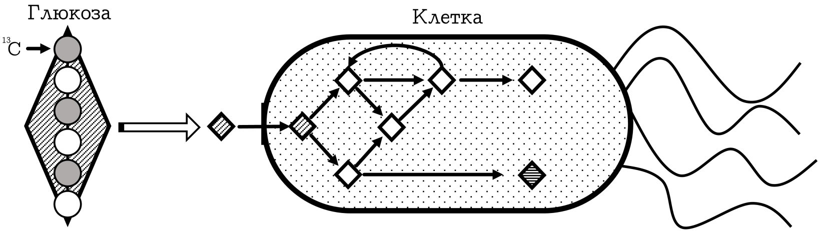
Анализ Метаболических Потоков

Метаболический поток — внутриклеточная химическая реакция
^{13}*C-Metabolic Flux Analysis* (^{13}*C-MFA*) — метод измерения метаболических потоков

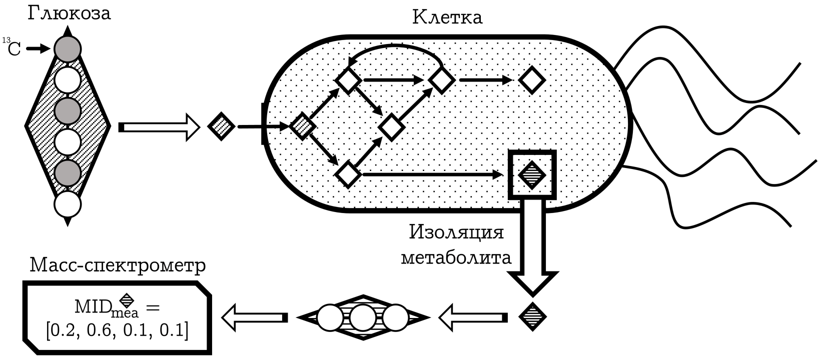
Анализ Метаболических Потоков



Анализ Метаболических Потоков



Анализ Метаболических Потоков



Анализ Метаболических Потоков

$$f(v) = \underset{\text{calc}}{\blacklozenge}$$

$$\arg_v f(v) \approx \underset{\text{mea}}{\blacklozenge}$$

Анализ Метаболических Потоков

- $\min_{\mathbf{v} \in U} \left(\mathbf{x}_{mea} - \mathbf{x}_{calc}(\mathbf{v})\right)^T \times \boldsymbol{\Sigma}^{-1} \times \left(\mathbf{x}_{mea} - \mathbf{x}_{calc}(\mathbf{v})\right)$
- Линейное программирование
- Метод оптимизации
- Конструирование графов
- Создание СЛАУ
- Решение СЛАУ
- Статистика
- Кластеризация результатов

Анализ Метаболических Потоков

- $\min_{\mathbf{v}\in U}\left(\mathbf{x}_{mea}-\mathbf{x}_{calc}(\mathbf{v})\right)^{T}\times\mathbf{\Sigma}^{-1}\times\left(\mathbf{x}_{mea}-\mathbf{x}_{calc}(\mathbf{v})\right)$
- Линейное программирование
- Метод оптимизации
- Конструирование графов
- Создание СЛАУ
- Решение СЛАУ
- Статистика
- Кластеризация результатов

Анализ Метаболических Потоков

- $\min_{\mathbf{v}\in U}\left(\mathbf{x}_{mea}-\mathbf{x}_{calc}(\mathbf{v})\right)^T \times \mathbf{\Sigma}^{-1} \times \left(\mathbf{x}_{mea}-\mathbf{x}_{calc}(\mathbf{v})\right)$
- Линейное программирование
- Метод оптимизации
- Конструирование графов
- Создание СЛАУ
- Решение СЛАУ
- Статистика
- Кластеризация результатов

Анализ Метаболических Потоков

- $\min_{\mathbf{v} \in U} \left(\mathbf{x}_{mea} - \mathbf{x}_{calc}(\mathbf{v})\right)^T \times \mathbf{\Sigma}^{-1} \times \left(\mathbf{x}_{mea} - \mathbf{x}_{calc}(\mathbf{v})\right)$
- Линейное программирование
- Метод оптимизации
- Конструирование графов
- Создание СЛАУ
- Решение СЛАУ
- Статистика
- Кластеризация результатов

Анализ Метаболических Потоков

- $\min_{\mathbf{v}\in U} \left(\mathbf{x}_{mea} - \mathbf{x}_{calc}(\mathbf{v})\right)^T \times \mathbf{\Sigma}^{-1} \times \left(\mathbf{x}_{mea} - \mathbf{x}_{calc}(\mathbf{v})\right)$
- Линейное программирование
- Метод оптимизации
- Конструирование графов
- Создание СЛАУ
- Решение СЛАУ
- Статистика
- Кластеризация результатов

Анализ Метаболических Потоков

- $\min_{\mathbf{v}\in U}\left(\mathbf{x}_{mea}-\mathbf{x}_{calc}(\mathbf{v})\right)^T \times \mathbf{\Sigma}^{-1} \times \left(\mathbf{x}_{mea}-\mathbf{x}_{calc}(\mathbf{v})\right)$
- Линейное программирование
- Метод оптимизации
- Конструирование графов
- Создание СЛАУ
- Решение СЛАУ
- Статистика
- Кластеризация результатов

Анализ Метаболических Потоков

- $\min_{\mathbf{v}\in U} \left(\mathbf{x}_{mea} - \mathbf{x}_{calc}(\mathbf{v})\right)^T \times \mathbf{\Sigma}^{-1} \times \left(\mathbf{x}_{mea} - \mathbf{x}_{calc}(\mathbf{v})\right)$
- Линейное программирование
- Метод оптимизации
- Конструирование графов
- Создание СЛАУ
- Решение СЛАУ
- Статистика
- Кластеризация результатов

Анализ Метаболических Потоков

- $\min_{\mathbf{v}\in U} \left(\mathbf{x}_{mea} - \mathbf{x}_{calc}(\mathbf{v})\right)^T \times \mathbf{\Sigma}^{-1} \times \left(\mathbf{x}_{mea} - \mathbf{x}_{calc}(\mathbf{v})\right)$
- Линейное программирование
- Метод оптимизации
- Конструирование графов
- Создание СЛАУ
- Решение СЛАУ
- Статистика
- Кластеризация результатов

Программы для ^{13}C-MFA расчетов

- 13CFLUX2
- Metran
- OpenFlux(2)
- FluxPyt
- mfapy
- Sysmetab
- baMFA
- iso2flux
- Flux-P
- WUFlux
- OpenMebius
- influx_s

Постановка задачи

- Написать эффективную программу для расчета ^{13}C-MFA на языке C++
- Провести тестирование, сравнить скорость работы с существующими аналогами

Программа Khnum

Используемые библиотеки:

- Eigen
- Alglib
- glpk
- Catch2

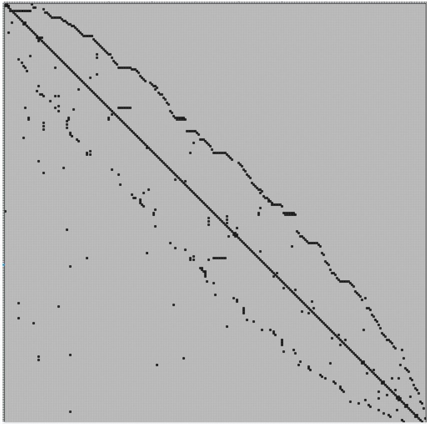
https://github.com/SteshinSS/khnum

Замеры

Метаболическая модель из 169 реакций
OpenFlux — 35 минут
Khnum, один поток — 22 секунды
Khnum, шестнадцать потоков — 4 секунды

Матрицы метода

$$\begin{bmatrix} -v_4 & v_4 & 0 & 0 & 0 \\ 0 & -v_1 - v_3 & v_3 & 0 & 0 \\ 0 & v_2 & -v_2 - v_5 & v_5 & 0 \\ 0 & 0 & 0 & -v_1 - v_3 & v_3 \\ v_5 & 0 & 0 & v_2 & -v_2 - v_5 \end{bmatrix}$$



М-матрицы

Определение М-матрицы (Ostrowsky, 1937)

Квадратная матрица $\mathbf{A} \in \mathbb{R}^{n \times n}$ называется *М-матрицей*, если:

1. Ее диагональные элементы больше или равны нулю $a_{ii} \geq 0$, $i = j$
2. Ее внедиагональные элементы меньше или равны нулю $a_{ij} \leq 0$, $i \neq j$
3. Матрицу $\mathbf{A}$ можно представить в виде: $\mathbf{A} = s\mathbf{I} - \mathbf{B}$, где $s > 0$, $\mathbf{B} \geq 0$, $s > \rho(\mathbf{B})$, где $\rho(\mathbf{B})$ — спектральный радиус $\mathbf{B}$

М-матрицы

Критерий М-матриц (Fiedler, Ptak, 1962)

Квадратная матрица является М-матрицей тогда и только тогда, когда она невырожденная и все вещественные собственные значения ее главных миноров больше или равны нулю.

М-матрицы

Теорема кругов (Гершгорин, 1931)

Пусть $\mathbf{A} \in \mathbb{C}^{n\times n}$ — комплексная матрица. Пусть $R_i = \sum_{i\neq j} |a_{ij}|$ — сумма модулей внедиагональных элементов i строки. Кругом Гершгорина назовем замкнутый круг $D(a_{ii}, R_i)$ с центром в a_{ii} и радиусом R_i. Тогда каждое собственное значение матрицы $\mathbf{A}$ лежит хотя бы в одном круге Гершгорина.

М-матрицы

Теорема-результат

Матрица коэффициентов MFA-систем, взятая со знаком минус, является М-матрицей.

ILU-разложение

Определение ILU-разложения (Meijerink, van der Vorst, 1977)

Пусть $\mathbf{A} \in \mathbb{R}^{n \times n}$ — разреженная матрица. Определим для нее *разреженную структуру* $S = \{(i,j) | a_{ij} \neq 0\} \cup \{(i,i)\}$ состоящую из всех координат ненулевых элементов и всех диагональных координат. Назовем *ILU-разложением* разложение вида $\mathbf{A} = \mathbf{LU} - \mathbf{R}$, где

- $\mathbf{L} \in \mathbb{R}^{n \times n}$ — нижнетреугольная матрица
- $\mathbf{U} \in \mathbb{R}^{n \times n}$ — верхнетреугольная матрица
- $\mathbf{L}, \mathbf{U}$ равны нулю вне разреженной структуры: $\mathbf{L}_{ij} = \mathbf{U}_{ij} = 0 \forall (i,j) \notin S$
- $\mathbf{R} \in \mathbb{R}^{n \times n}$ равна нулю в разреженной структуре: $\mathbf{R}_{ij} = 0 \forall (i,j) \in S$

Замеры

М-матрица 253×253

- BiCGSTAB + ILU
- BiCGSTAB + Diag
- LU (Partial Pivoting)
- SuperLU + COLAMD

Таблица: Сравнение методов. Время в микросекундах

BiCGSTAB + ILU	BiCGSTAB + Diag	DenseLU	SparseLU
125	986	645	174

Продуктивная матрица

Определение продуктивной матрицы (Леонтьев, 1928)

Квадратная вещественная матрица $\mathbf{A} \in \mathbb{R}^{n \times n}$ называется продуктивной, если

- $a_{ij} \leq 0, \forall i \neq j$
- $\exists x \geq 0 : Ax > 0$

Сравнение моделей

- Есть несколько отраслей
- Они используют ресурсы друг друга, чтобы производить ресурсы
- Экономика сильна

- Есть несколько химических реакций
- Они используют метаболиты друг друга, чтобы производить метаболиты
- Клетка жива

Ряд Неймана

Обращение M-матрицы (F. Waugh, 1950)

Пусть $\mathbf{A} = c\mathbf{I} - \mathbf{B}$ — невырожденная M-матрица. Тогда $\mathbf{A} = \sum_{k=0}^{\infty} \frac{\mathbf{A}^k}{c^{k+1}}$, причем ряд абсолютно сходится.

Temporal Leontief Inverse

M. Sonis, G. Hewings, 1998

$\mathbf{A}_0\mathbf{x} = \mathbf{y}$

$\mathbf{x} = \mathbf{A}_0^{-1}\mathbf{y}$

$\mathbf{A}_0^{-1} = \mathbf{R}$

$\mathbf{A}_1 = \mathbf{A}_0 + \mathbf{E}$

$\mathbf{A}_1^{-1} = \mathbf{R}(\mathbf{E})$

$\mathbf{R}(\mathbf{E})-?$

Temporal Leontief Inverse

M. Sonis, G. Hewings, 1998

$\mathbf{A}_0\mathbf{x} = \mathbf{y}$

$\mathbf{x} = \mathbf{A}_0^{-1}\mathbf{y}$

$\mathbf{A}_0^{-1} = \mathbf{R}$

$\mathbf{A}_1 = \mathbf{A}_0 + \mathbf{E}$

$\mathbf{A}_1^{-1} = \mathbf{R}(\mathbf{E})$

$\mathbf{R}(\mathbf{E})$—?

Temporal Leontief Inverse

M. Sonis, G. Hewings, 1998

$\mathbf{A}_0\mathbf{x} = \mathbf{y}$

$\mathbf{x} = \mathbf{A}_0^{-1}\mathbf{y}$

$\mathbf{A}_0^{-1} = \mathbf{R}$

$\mathbf{A}_1 = \mathbf{A}_0 + \mathbf{E}$

$\mathbf{A}_1^{-1} = \mathbf{R}(\mathbf{E})$

$\mathbf{R}(\mathbf{E})$–?

Temporal Leontief Inverse

M. Sonis, G. Hewings, 1998

$\mathbf{A}_0\mathbf{x} = \mathbf{y}$

$\mathbf{x} = \mathbf{A}_0^{-1}\mathbf{y}$

$\mathbf{A}_0^{-1} = \mathbf{R}$

$\mathbf{A}_1 = \mathbf{A}_0 + \mathbf{E}$

$\mathbf{A}_1^{-1} = \mathbf{R}(\mathbf{E})$

$\mathbf{R}(\mathbf{E})$—?

Temporal Leontief Inverse

M. Sonis, G. Hewings, 1998

$\mathbf{A}_0\mathbf{x} = \mathbf{y}$

$\mathbf{x} = \mathbf{A}_0^{-1}\mathbf{y}$

$\mathbf{A}_0^{-1} = \mathbf{R}$

$\mathbf{A}_1 = \mathbf{A}_0 + \mathbf{E}$

$\mathbf{A}_1^{-1} = \mathbf{R}(\mathbf{E})$

$\mathbf{R}(\mathbf{E})-?$

Temporal Leontief Inverse

M. Sonis, G. Hewings, 1998

$\mathbf{A}_0\mathbf{x} = \mathbf{y}$

$\mathbf{x} = \mathbf{A}_0^{-1}\mathbf{y}$

$\mathbf{A}_0^{-1} = \mathbf{R}$

$\mathbf{A}_1 = \mathbf{A}_0 + \mathbf{E}$

$\mathbf{A}_1^{-1} = \mathbf{R}(\mathbf{E})$

$\mathbf{R}(\mathbf{E})-?$

Обратная для симметричной диагонально доминантной М-матрицы

$$\mathbf{A}\mathbf{x} = \mathbf{y}$$

$$\mathbf{A}^{\mathsf{T}}\mathbf{A}\mathbf{x} = \mathbf{A}^{\mathsf{T}}\mathbf{y}$$

$$\mathbf{x} = (\mathbf{A}^{\mathsf{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathsf{T}}\mathbf{y}$$

Обратная для симметричной диагонально доминантной M-матрицы

$$\mathbf{A}\mathbf{x} = \mathbf{y}$$

$$\mathbf{A}^{\mathsf{T}}\mathbf{A}\mathbf{x} = \mathbf{A}^{\mathsf{T}}\mathbf{y}$$

$$\mathbf{x} = (\mathbf{A}^{\mathsf{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathsf{T}}\mathbf{y}$$

Обратная для симметричной диагонально доминантной M-матрицы

$$\mathbf{A}\mathbf{x} = \mathbf{y}$$

$$\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{y}$$

$$\mathbf{x} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{y}$$

Обратная для симметричной диагонально доминантной M-матрицы

Теорема (C. Dellacherie, S. Martinez, J.S. Matrin, 2014)

$$\mathbf{A} = \begin{bmatrix} x+a+b & -a & -b \\ -a & y+a+c & -c \\ -b & -c & z+b+c \end{bmatrix}$$

$$\mathbf{A}^{-1} = \frac{1}{\det} \begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{bmatrix}^{\mathsf{T}}$$

$$\det = (x+a+b)(y+a+c)(z+b+c)$$

Обратная для симметричной диагонально доминантной M-матрицы

Теорема (C. Dellacherie, S. Martinez, J.S. Matrin, 2014)

$$C_{11} = (y+a+c)(z+b+c) = yz + by + cy + az + ab + ac + cz + bc$$

$$C_{22} = (x+a+b)(z+b+c) = xz + bx + cx + az + ab + ac + bz + bc$$

$$C_{33} = (x+a+b)(y+a+c) = xy + ax + cx + ay + ac + by + ab + bc$$

$$C_{13} = C_{31} = C_{11} \cap C_{33} = by + ab + ac + bc$$

$$C_{12} = C_{21} = C_{11} \cap C_{22} = az + ab + ac + bc$$

$$C_{23} = C_{32} = C_{22} \cap C_{33} = cx + ab + ac + bc$$

Обратная для симметричной диагонально доминантной M-матрицы

Теорема (C. Dellacherie, S. Martinez, J.S. Matrin, 2014)

$$C_{11} = (y + a + c)(z + b + c) = yz + by + cy + az + ab + ac + cz + bc$$
$$C_{22} = (x + a + b)(z + b + c) = xz + bx + cx + az + ab + ac + bz + bc$$
$$C_{33} = (x + a + b)(y + a + c) = xy + ax + cx + ay + ac + by + ab + bc$$

$$C_{13} = C_{31} = C_{11} \cap C_{33} = by + ab + ac + bc$$
$$C_{12} = C_{21} = C_{11} \cap C_{22} = az + ab + ac + bc$$
$$C_{23} = C_{32} = C_{22} \cap C_{33} = cx + ab + ac + bc$$

Обратная для симметричной диагонально доминантной M-матрицы

Теорема (C. Dellacherie, S. Martinez, J.S. Matrin, 2014)

$$\mathbf{A} = \begin{bmatrix} x+a+b & -a & -b \\ -a & y+a+c & -c \\ -b & -c & z+b+c \end{bmatrix}$$

$$\mathbf{A}^{-1} = \frac{1}{\det} \begin{bmatrix} (y+a+c)(z+b+c) & az+t & by+t \\ az+t & (x+a+b)(z+b+c) & cx+t \\ by+t & cx+t & (x+a+b)(y+a+c) \end{bmatrix}^{\mathsf{T}}$$

$$\det = (x+a+b)(y+a+c)(z+b+c)$$

$$t = ab+ac+bc$$

Полученные результаты

- Написана эффективная открытая программа Khnum для проведения ^{13}C-MFA расчетов.
- Проведено сравнение с аналогами.
- Доказана принадлежность матрицы коэффициентов СЛАУ к классу M-матриц. Это позволило использовать специальный предобуславливатель на основе ILU-разложения.
- Проведено сравнение нескольких численных методов для СЛАУ Показано, что численные методы с ILU-предобуславливателем работают быстрее всего.

Спасибо за внимание