

Intracellular Metabolic Networks and Metabolic Flux Analysis

Jeffrey D. Varner*

School of Chemical Engineering*

Purdue University, West Lafayette IN 47907

Copyright © Jeffrey Varner 2016. All Rights Reserved.

Introduction

Cells operate vast interconnected biochemical *networks* to process starting materials such as sugars into valuable products of interest, waste products and eventually more cells. Previously, we derived a material balance governing the specific intracellular concentration of the j^{th} metabolite x_j [mmol/gdw] in an arbitrary metabolic network:

$$\frac{dx_j}{dt} = \left(\sum_r \sigma_{jr} v_r + \sum_k \tau_{jk} q_k \right) - \mu x_j \quad j = 1, 2, \dots, \mathcal{M} \quad (1)$$

or in matrix-vector:

$$\frac{d\mathbf{x}}{dt} = \mathbf{S}\mathbf{v} + \mathbf{T}\mathbf{q} - \mu\mathbf{x} \quad (2)$$

where \mathbf{x} denotes the species abundance vector ($\mathcal{M} \times 1$), \mathbf{S} denotes the stoichiometric matrix ($\mathcal{M} \times \mathcal{R}$), and \mathbf{T} ($\mathcal{M} \times \mathcal{T}$) denotes the transport (or exchange) matrix. Given these balance equations, we could estimate the steady-state flux (specific reaction rates \mathbf{v} , \mathbf{q}) through the network using flux balance analysis:

$$\begin{aligned} & \max_{\mathbf{v}} (v_{obj} = \mathbf{c}^T \mathbf{v}) \\ & \text{Subject to : } \mathbf{S}\mathbf{v} = -\mathbf{T}\mathbf{q} \\ & \alpha_i \leq v_i \leq \beta_i \end{aligned} \quad (3)$$

where \mathbf{c}^T is the *objective selection vector* and \mathbf{S} and \mathbf{T} denote the stoichiometric and transport arrays for both intracellular and extracellular metabolites (rows) for unmeasured and measured fluxes (columns). However, flux balance analysis is not the only technique that can be used to estimate intracellular metabolic flux. Metabolic flux analysis (MFA) is another widely used approach that relies on linear algebra (and not linear programming) to compute the unknown metabolic flux vector.

Metabolic Flux Analysis (MFA). Similar to flux balance analysis, metabolic flux analysis relies upon a pseudo steady-state assumption to reduce the intracellular material balance equations to a system of algebraic equations:

$$\mathbf{S}\mathbf{v} + \mathbf{T}\mathbf{q} - \mu\mathbf{x}^* \simeq \mathbf{0} \quad (4)$$

where \mathbf{x}^* denotes the steady-state species abundance vector. Eqn (4) has two sets of unknowns, the states and the fluxes. Without direct experimental measurements of the intracellular state of the cell, we have way of knowing \mathbf{x}^* . Moreover, we know the dilution due to growth terms ($\mu\mathbf{x}^*$) are small compared to the metabolic flux terms. Thus, we typically drop the dilution terms to arrive at:

$$\mathbf{S}\mathbf{v} + \mathbf{T}\mathbf{q} \simeq \mathbf{0} \quad (5)$$

Eqn (5) are \mathcal{M} algebraic equations in the unknown intracellular fluxes $v_1, \dots, v_{\mathcal{R}}$ and transport fluxes $q_1, \dots, q_{\mathcal{T}}$. Typically, $\mathcal{M} < \mathcal{R} + \mathcal{T}$ which means that no unique flux solution can be found. However, we can potentially measure some of the transport fluxes that carry material into and from the cells. Denote the subset of *measured* fluxes as ϑ_m ($\mathcal{E} \times 1$ column vector), and all other unmeasured fluxes as ϑ_u ($\mathcal{U} \times 1$ column vector). We can then redefine the stoichiometric balances as:

$$\Sigma\vartheta_u + \Phi\vartheta_m \simeq \mathbf{0} \quad (6)$$

where Φ contains the columns corresponding to the measured fluxes ($\mathcal{M} \times \mathcal{E}$), and Σ contains the columns corresponding to *unmeasured* fluxes ($\mathcal{M} \times \mathcal{U}$). Eqn (6) can be solved for the *unmeasured* fluxes ϑ_u :

$$\vartheta_u \simeq -\Sigma^\# \Phi \vartheta_m \quad (7)$$

where $\Sigma^\#$ is a $\mathcal{U} \times \mathcal{M}$ *generalized* matrix inverse. The functional form of the generalized inverse depends upon the shape and properties of the unmeasured stoichiometric array Σ .

The matrix $\Sigma^\#$ is the generalized inverse of the stoichiometric matrix corresponding to the *unmeasured* fluxes. The form of $\Sigma^\#$ depends upon the shape and the *rank* of Σ . The definition of the Σ and Φ arrays influences our ability to solve the intracellular material balances. The structure of these arrays is controlled practically by which fluxes are measured. Measurement selection is influenced by several factors, technical concerns regarding the measurement technology, the cost of measurements as well as the time required to make measurements.

- $\mathcal{M} > \mathcal{U}$: **Overdetermined system of equations.** Overdetermined systems are characterized by more equations than unknowns and have no single *regular* solution, in other words, no single solution that simultaneously satisfies all material balance equations. The problem of determining the solution to an overdetermined system can be recast as the least-squares problem:

$$\min_{\vartheta_u} (\Sigma\vartheta_u - \mathbf{b})^T (\Sigma\vartheta_u - \mathbf{b}) \quad (8)$$

where

$$\mathbf{b} \equiv -\Phi\vartheta_m(t) \quad (9)$$

Equation (8) can be solved analytically by using standard tools from calculus to yield the solution:

$$\vartheta_u = (\Sigma^T \Sigma)^{-1} \Sigma^T \mathbf{b} \quad (10)$$

In this case the quantity $\Sigma^\#$ is given by:

$$\Sigma^\# = \Sigma^L \equiv (\Sigma^T \Sigma)^{-1} \Sigma^T \quad (11)$$

denotes a *left* inverse (sometimes called the left Moore-Penrose inverse, left generalized inverse or pseudo inverse). For the least-squares solution Equation (10) to exist, the condition

$$\det \Sigma^T \Sigma \neq 0 \quad (12)$$

must be satisfied.

- $\mathcal{M} < \mathcal{U}$: **Underdetermined system of equations.** Underdetermined systems are characterized by ∞ -solutions as n unknowns can be written in terms of the remaining $m - n$ unknowns which in the general case take on arbitrary values. The solution of an underdetermined system of LAEs can be posed as a constrained least-squares minimization problem of the form:

$$\min_{\mathbf{u}} \frac{1}{2} \vartheta_{\mathbf{u}}^T \vartheta_{\mathbf{u}} \quad (13)$$

subject to:

$$\Sigma \vartheta_{\mathbf{u}} = \mathbf{b} \quad (14)$$

where

$$\mathbf{b} \equiv -\Phi \vartheta_{\mathbf{m}}(t) \quad (15)$$

Problem (13) can also be solved analytically using standard tools from calculus to yield the solution:

$$\vartheta_{\mathbf{u}} = \Sigma^T (\Sigma \Sigma^T)^{-1} \mathbf{b} \quad (16)$$

The quantity

$$\Sigma^\# = \Sigma^R \equiv \Sigma^T (\Sigma \Sigma^T)^{-1} \quad (17)$$

denotes a *right* inverse (sometimes called the right Moore-Penrose inverse, right generalized inverse or pseudo inverse). For the least-squares solution Equation (17) to exist, the condition

$$\det \Sigma \Sigma^T \neq 0 \quad (18)$$

must be satisfied.

- $\mathcal{M} = \mathcal{U}$: **Square system of equations.** When the number of equations equals the number of unknowns the system is said to be *square*. Square systems can be solved by several different techniques, including directly determining the matrix inverse Σ^{-1} , i.e.,

$$\vartheta_{\mathbf{u}} = -\Sigma^{-1} \Phi \vartheta_{\mathbf{m}}(t) \quad (19)$$

Solution (19) exists as long as:

$$\det(\Sigma) \neq 0 \quad (20)$$

However, the direct computation of the inverse for square systems is *very* computationally expensive. Thus, in practice other techniques, such as Gauss Elimination (GE), Gauss Jordan Elimination (GJE) or iterative techniques (ITs) are used to calculate the flux vector v_u .

An alternative to computing the determinant to check the existence of the matrix inverse, is to compute the rank. The rank of the $N \times M$ matrix \mathbf{A} given by:

$$\text{rank}(\mathbf{A}) \leq \min(N, M) \quad (21)$$

is a measure of the independent information contained in the rows and columns of a matrix. Arrays with linearly independent rows and columns have full rank, while arrays with linearly dependent components are not full rank. For a determinant to be non-zero, an array has to be full rank. Performing a rank test is more efficient than computing the determinant of a matrix, especially for large non-square matrices.

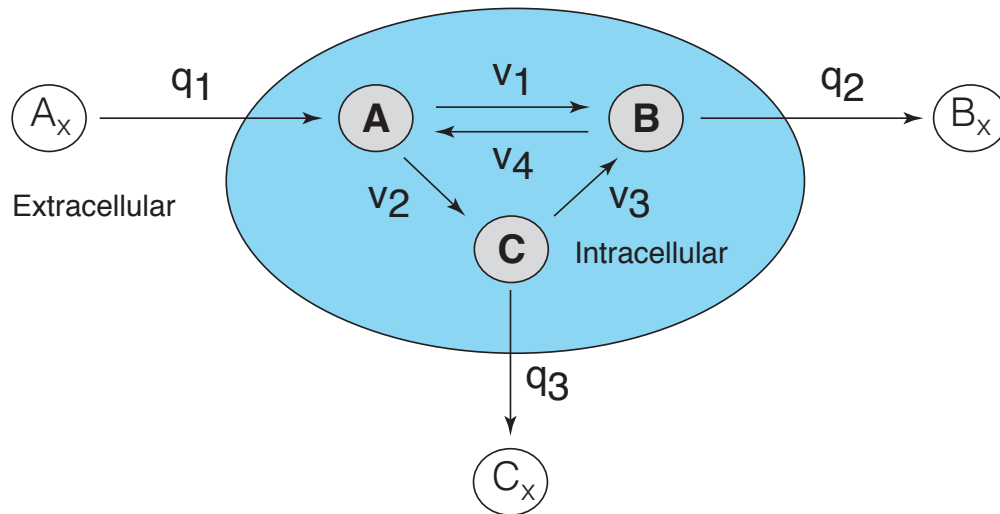


Fig. 1: Schematic of simple metabolic reaction network. This system has a total of six metabolites and seven reactions. The intracellular metabolites (metabolites A , B and C) are often treated as balanced (no accumulation) while the extracellular metabolites (A_x , B_x and C_x) are not balanced (can accumulate). The blue-oval denotes the cell boundaries, where q_j denotes flux across these boundaries ([mmol/gdw-hr]) and v_k denotes the intracellular fluxes.

Example: Proof of concept MFA with a toy network. Estimate the intracellular fluxes for the simple reaction network shown in Fig. (1) using metabolic flux analysis (MFA). Assuming we measured all three transport fluxes, the stoichiometric and transport arrays for this network are given by:

$$\Sigma = \begin{bmatrix} -1 & -1 & 0 & 1 \\ 1 & 0 & 1 & -1 \\ 0 & 1 & -1 & 0 \end{bmatrix} \quad (22)$$

and

$$\Phi = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad (23)$$

The stoichiometric matrix Σ is not square nor does it have full rank $\text{rank}(\Sigma) = 2$. Thus, even the pseudo inverse is not defined for this set of measured fluxes, suggesting we need to change our experimental design. If we measure only q_1 and q_2 , the stoichiometric and transport arrays become:

$$\Sigma = \begin{bmatrix} -1 & -1 & 0 & 1 & 0 \\ 1 & 0 & 1 & -1 & 0 \\ 0 & 1 & -1 & 0 & -1 \end{bmatrix} \quad (24)$$

and

$$\Phi = \begin{bmatrix} 1 & 0 \\ 0 & -1 \\ 0 & 0 \end{bmatrix} \quad (25)$$

which has full rank $\text{rank}(\Sigma) = 3$. We can compute a right inverse and solve for $\vartheta_u = (v_1, v_2, v_3, q_3)$. If we assume $q_1 = 1.0$ and $q_2 = 0.5$, the flux solution is shown in Fig. (2).

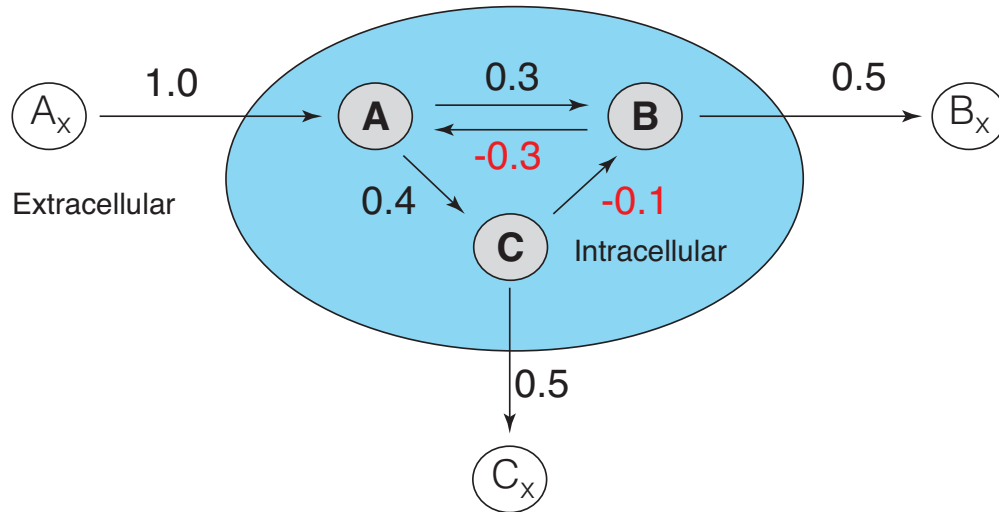


Fig. 2: Example flux solution for $q_1 = 1.0$ and $q_2 = 0.5$. While the overall material balance is consistent, some estimated intracellular fluxes (v_3 and v_4) violate non-negativity.

We can resolve the non-negativity issue by further constraining the relationships between the fluxes. However, the formulation of these additional constraints are case specific. In the simplest case, let's assume there is no flux through v_4 . Thus, we move v_4 to the measured list and and redefine Σ and Φ as:

$$\Sigma = \begin{bmatrix} -1 & -1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & -1 \end{bmatrix} \quad (26)$$

and

$$\Phi = \begin{bmatrix} 1 & 0 & 1 \\ 0 & -1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \quad (27)$$

The $\text{rank}(\Sigma) = 3$, and the solution (assuming $q_1 = 1.0$, $q_2 = 0.5$ and $v_4 = 0.0$) is shown in Fig. (3).

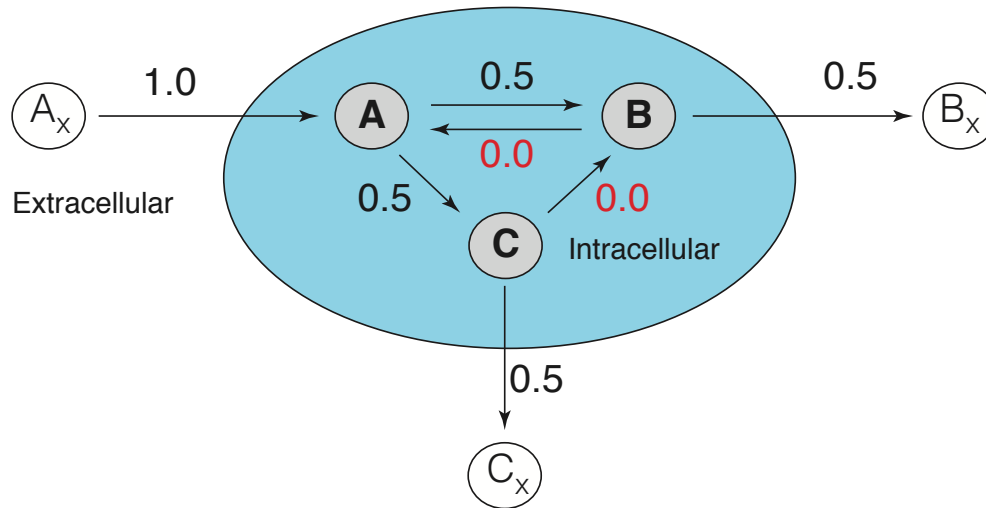


Fig. 3: Example flux solution for $q_1 = 1.0$, $q_2 = 0.5$ and $v_4 = 0.0$. The overall material balance is consistent, and no intracellular fluxes violate non-negativity. However, we required an additional measurement (constraint) to calculate the solution