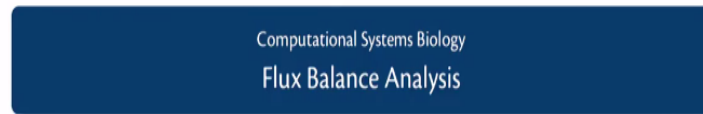


Computational Systems Biology
Karthik Raman
Department of Biotechnology
Indian Institute of Technology – Madras

Lecture - 57
Flux Balance Analysis

(Refer Slide Time: 00:11)



- ▶ Steady-state Mass Balance
- ▶ Constraints
- ▶ Optimisation (LP)

Karthik Raman

Department of Biotechnology, Bhupar & Jyoti Mehta School of Biosciences
Initiative for Biological Systems Engineering (IBSE)
Robert Bosch Centre for Data Science and Artificial Intelligence (RBC DSAI)
INDIAN INSTITUTE OF TECHNOLOGY MADRAS



In this video, I will introduce you to Flux Balance Analysis wherein we will see how we can take the equation from the previous class and setup a steady state mass balance problem and impose further constraints on it and finally optimize to reliably predict the growth rate of a cell or any biochemical reaction network.

(Refer Slide Time: 00:31)



FLUX BALANCE ANALYSIS (FBA)



So flux balance analysis is the most important concept when it comes to constraint based modeling, it is one of the first take things that was you know used.

(Refer Slide Time: 00:45)

Flux balance analysis (FBA)

Bonarius HPI et al. (1997) *Trends in Biotechnology* 15:308-314

► Steady state mass balance ($S \cdot v = 0$)

$$\text{► } v = [v_1 \ v_2 \ \dots \ v_n \ b_1 \ b_2 \ \dots \ b_{\text{out}}]^T$$

► v_i are internal fluxes

► Addition of Constraints:

► Optimisation: Maximise growth (or) Minimise nutrient uptake (or) Maximise metabolite production (or) Euclidean norm (efficient metabolite channelling)

► Mathematically a linear programming (LP) problem: $\min, c^T v \quad \text{s.t. } S \cdot v = 0$

► Perturbation analysis in silico: gene deletion, drug inhibition



So it involves a steady state mass balance, right. What is steady state? There is no accumulation, right. So whatever substance is coming out there are some conversions and steady state substances are going out, right. So what would be dx/dt under steady state?

(Refer Slide Time: 01:08)

Stoichiometric matrix
 $S_{m \times r}$
 $m > r$?
 $m < r$?
 $m = r$?

$\vec{0} = \frac{d\vec{v}}{dt} = S \cdot \vec{v} = 0$
 metabolite concentrations
 unknowns (fluxes)

$v_1 > 0$
 $v_2 > 0$
 $v_1 = v_2$
 $\& v_4 < 5$

It will be = 0. The rate of change of every rate of accumulation of every metabolite will be 0. Is that right? So what does that mean it means that $S \cdot v = 0$.

(Refer Slide Time: 01:35)

$S \cdot v = 0 \Rightarrow ?$
 System of linear eqs
 under-determined system $m < r$
 over-determined system $m > r$
 'exact' system $m = r$
 Metabolic networks

$m = \# \text{ metabolites} \equiv \# \text{ eqs}$
 $n = \# \text{ rows / fluxes} \equiv \# \text{ unknowns}$

$v_1 + v_2 = 10$

What is this mean? What is this? What does $S \cdot v = 0$ mean? It is a system of linear equations. Is it-- the moment you have a system of linear equations you could have an under-determined system and an over - or a you know r you could call an exact system, right. So this is $m=r$; what does over-determined mean? $m>r$, $m<r$. So once again what is m ? which is the same as the number of equations. Every metabolite has an equation.

R, number of reactions of fluxes which is a same as variables or unknowns. So if I just gave you an equation $v_1+v_2=10$ and ask you to solve it, how many solutions would you have? Obviously, infinitely many solutions, which is the same thing we can expect for any underdetermined system. Whenever you have more reactions than metabolites you will have an undetermined system or a fat matrix, right.

So when you have an underdetermined system you now need some other ways of solving it. What happens if you have exact system? **“Professor – student conversation starts”** (()) (04:07) And what is that solution? All 0s, right hand side is 0, right? So you know that all 0s is a solution. You know that it is therefore the solution, right. **“Professor – student conversation ends”**

And over determined system it could be inconsistent, right or you might be able to fit a good solution, a solution that satisfies most constraints within some epsilons something you can do early least square solution and so on. So let us focus mostly on underdetermined systems, which is always we always observed in biological system or metabolic networks, right. So now you have a system $Sv=0$, how do you solve it? So let us just recap.

(Refer Slide Time: 05:17)

$Sv = 0 \Rightarrow ?$
 System of linear eqs:
 - under-determined system $m < r$
 - over-determined system $m > r$
 - 'exact' system $m = r$

$m = \# \text{ metabolites} \equiv \# \text{ eqs}$
 $n = \# \text{ rxns / fluxes} \equiv \# \text{ unknowns}$

Metabolic network:
 $Sv = 0$
 $S_{m \times n} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$
 $v_1 + v_2 = 10$

$\phi \rightarrow A$
 $A_{\text{node}} \rightarrow A$
 $D \rightarrow -1$

So we are looking at a steady state mass balance of $Sv=0$, right. And we want to look at what is v , so v is the set of fluxes, I write it in a horizontal fashion of the transpose so save space on the

slide. So v is nothing but, and you can also think of external fluxes, not be an n , it can be an m , right. This is going to be your r cross one vector that you are going to find out. This is what multiplies s . What is the set of this vector? Right, this is going to be m cross, this is m cross 1, this is m cross 1. Clear?

So you basically discriminate between the internal fluxes and the external fluxes. External fluxes involve exchange with the outer environment. So you will typically have a cell like this. All these are internal fluxes, and let us say, so think of a cell, you have several internal fluxes internal reactions so these are metabolites and you can think of all these as internal fluxes and then you can also have some nutrients coming in, things being secreted outside by the cell, important products and so on.

So these are what we call the exchange fluxes and these are the internal fluxes and so on. **“Professor – student conversation starts”** (07:56) These are not necessarily reactions. So could still right them as reaction. I would you just right them as or essentially A coming in to the cell, okay. And this would just be something like d going out. **“Professor – student conversation ends”**

So this Stoichiometric matrix column corresponding to this would just be -1 and everything else is 0 . This would actually have been, you can also have it as A external giving A and so on. But the easiest way to write it has just you know A coming in, right so then you can balance it with the rest of it. It will become clearer then we look at in an example.

(Refer Slide Time: 09:04)

Flux balance analysis (FBA)

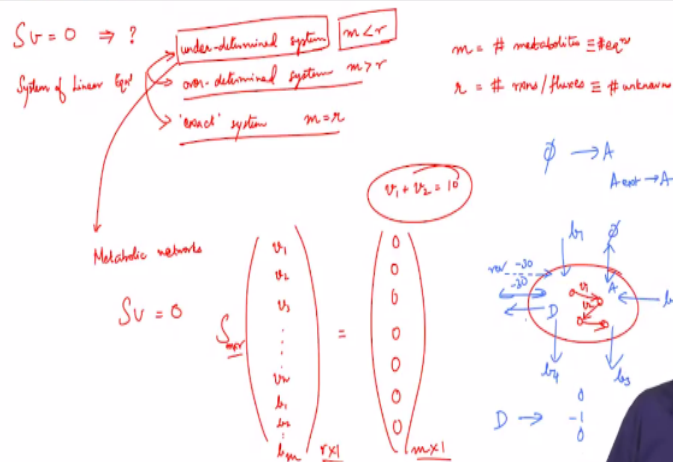
Bonarius HPI et al. (1997) Trends in Biotechnology 15:308-314

- ▶ Steady state mass balance ($S \cdot v = 0$)
- ▶ $v = [v_1 \ v_2 \ \dots \ v_m \ b_1 \ b_2 \ \dots \ b_{n_{\text{ext}}}]^T$
 - ▶ v_i are internal fluxes
- ▶ Addition of Constraints: ¹
 - ▶ Irreversible reactions: $0 \leq v_i < \infty$
 - ▶ Reversible reactions: $-\infty < v_j < \infty$
 - ▶ External fluxes/Uptake reactions/Secretions: $-\infty < b_i < \infty$
- ▶ Optimisation: Maximise growth (or) Minimise nutrient uptake (or) Maximise metabolite production (or) Euclidean norm (efficient metabolite channelling)
- ▶ Mathematically, a linear programming (LP) problem: $\min_v c^T v \quad \text{s.t.} \quad S \cdot v = 0$
- ▶ Perturbation analysis *in silico*: gene deletion, drug inhibition

And then we go and add constraints. What kind of constraints can we add? Stoichiometry is accounted for right, this is your Stoichiometry constraint right whereas you can add constraints on reversibility, irreversibility. So there are irreversible reactions which you will say can go only in one direction, so we can say $0 < v < \text{infinite}$, in practice no infinite but a large number. And you have reversible reactions which basically can go in either direction.

And you can have exchanges fluxes are Uptake that can either be reversible or irreversible. So could think of a particular flux being like this.

(Refer Slide Time: 09:57)

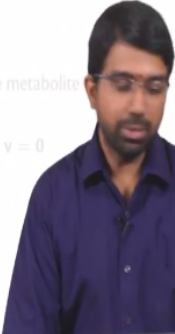


You can go in both directions. So now if you get this value as -30 it means it goes in this direction, you have to commit to an arrow sign in the first place but once you do that let us say this is a reversible reaction and if it now has the value -30 it means it is actually 30 moles per liter per liter pair are going in the opposite direction.

(Refer Slide Time: 10:21)

Flux balance analysis (FBA)
Bonarius HPI et al. (1997) Trends in Biotechnology 15:308-314

- ▶ Steady state mass balance ($S \cdot v = 0$)
- ▶ $v = [v_1 \ v_2 \ \dots \ v_{n_i} \ b_1 \ b_2 \ \dots \ b_{n_{ext}}]^T$
 - ▶ v_i are internal fluxes
- ▶ Addition of Constraints:
 - ▶ Irreversible reactions: $0 \leq v_i < \infty$
 - ▶ Reversible reactions: $-\infty < v_i < \infty$
 - ▶ External fluxes/Uptake reactions/Secretions: $-\infty < b_i < \infty$
- ▶ Optimisation: Maximise growth (or) Minimise nutrient uptake (or) Maximise metabolite production (or) Euclidean norm (efficient metabolite channelling)
- ▶ Mathematically, a linear programming (LP) problem: $\min_v c^T v \quad \text{s.t. } S \cdot v = 0$
- ▶ Perturbation analysis in silico: gene deletion, drug inhibition



So how do you go about solving this problem now? You have an over determined system which means you have to find out a single solution.

(Refer Slide Time: 10:34)

$Sv = 0 \Rightarrow ?$

System of linear eqs!

- under-determined system $m < r$
- over-determined system $m > r$
- 'exact' system $m = r$

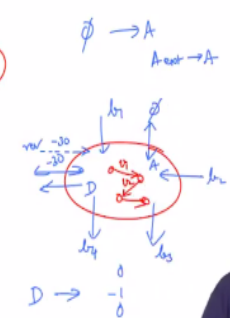
Metabolic network

$Sv = 0$

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_m \\ b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$m = \# \text{ metabolites} \equiv \# eqs$
 $n = \# rxns / fluxes \equiv \# unknowns$

$\max v_i \in \{0, \infty\}$
 $v_1 + v_2 = 10$



If you want to find out a single solution to the problem, you need to come up with some other constraints or some other strategy. How would do you go about it? So if I wanted to solve this,

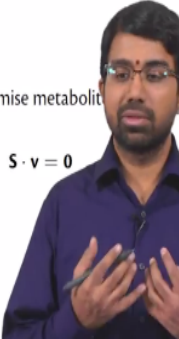
this system $v_1+v_2=10$ one way of potentially solving it is if I said told you some other conditions. If I said, this makes it a little better but the problem is still there I have no, I think this almost specifies a solution.

Now there is only one possible solution that is possible, so 5+5. Is that clear? So you need to give some other constraints or going for some optimization. So usually we resolve to maximizing something. Let us say solve this subject to max v_1 and the answer is very simple, 10,0. Solve this subject of max v_2 0, 10 right. So it is easy to solve it to find potentially a unique solution not necessarily always. When you start specifying in optimization criterion.

(Refer Slide Time: 12:19)

Flux balance analysis (FBA)
Bonarius HPJ et al. (1997) Trends in Biotechnology 15:308-314

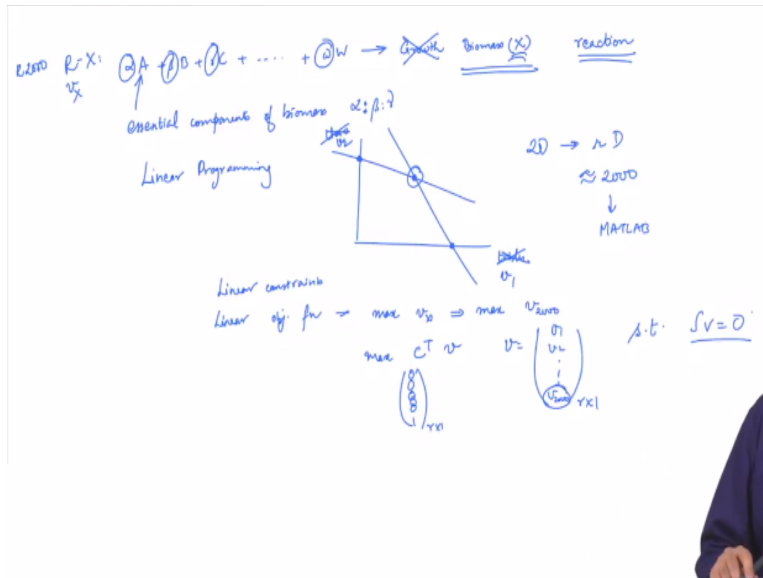
- ▶ Steady state mass balance ($S \cdot v = 0$)
- ▶ $v = [v_1 \ v_2 \ \dots \ v_{n_i} \ b_1 \ b_2 \ \dots \ b_{n_{ext}}]^T$
 - ▶ v_i are internal fluxes
- ▶ Addition of Constraints:
 - ▶ Irreversible reactions: $0 \leq v_i < \infty$
 - ▶ Reversible reactions: $-\infty < v_j < \infty$
 - ▶ External fluxes/Uptake reactions/Secretions: $-\infty < b_k < \infty$
- ▶ Optimisation: **Maximise growth** (or) Minimise nutrient uptake (or) Maximise metabolite production (or) Euclidean norm (efficient metabolite channelling)
- ▶ Mathematically, a linear programming (LP) problem: $\min_v \ c^T v \quad \text{s. t.} \quad S \cdot v = 0$
- ▶ Perturbation analysis in silico: gene deletion, drug inhibition



So typically, it is been found that if you try to maximum growth or minimize nutrient uptake they are useful optimizations to carry out to find out what is the possible flux distribution. So it turns out that maximizing growth is a very useful objective function. **“Professor – student conversation starts”** Sir, what set of (()) (12:47). So we will come to that, we basically say that maximize a growth flux.

So you have to potentially define flux that accounts for growth and then you try to maximize it. So what would that be, that is quite interesting. **“Professor – student conversation ends”**

(Refer Slide Time: 13:06)



So essentially set of reaction, let us first setup the reaction, in fact let us not call it growth we will call it Biomass since usually can be represented by the symbol x , fine. So let us say this is an additional reaction in your cell, fair enough. So how would, what would be needed for the cell to produce to produce this metabolite. We now call, this is a factious metabolite, Biomass is a factious metabolite x .

So it needs alpha moles of A, beta moles of B, gamma moles of B and omega moles of W to produce x , so all these can be your essential components of biomass in the correct ratio.

“Professor – student conversation starts” Have you solved this kind of a problem before, back in school? Linear programming. Linear programming, right, you are very familiar with.

“Professor – student conversation ends”

This is nothing but linear programming. And all of you must have potentially solved a problem where you had a factory that would make tables and chairs and there is only so many chairs they can make or so many tables that they can make and the profit on a table is the x and the profit on the chair is y , how many tables and how many chairs do you produce to maximize profit, right.

So you remember how you use to solve it, you would basically put it on a graph paper and then you will draw some lines corresponding to the different linear constraints that you had and then you would look at the corner points because that is where-- It much the same problem excepts

from 2D to move to a problem that is much larger, right it is r dimensions, right. And roughly 2000 dimension which means you need MATLAB or some other tool to solve it.

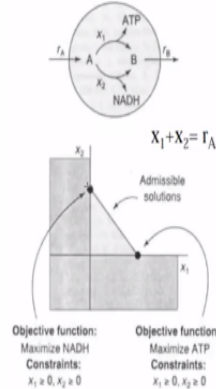
Good solver LP, Linear programming solver to solve it. But this is much the same problem as the linear programming you studied in school. Right, so this is mathematically, a linear programming problem, why is it a linear problem? Because you have linear constraints and most importantly a linear objective function, right. This function could be something like let us call this reaction x , right for Biomass production. So we would say the flux of this is some v_x .

So you say maximize v_x . Let us say for example this is the 2000th reaction. This is the same as saying maximize v_{2000} or maximize $C^T v$ you obviously know is v_1, v_2, v_{2000} and C is nothing but $0,0,0,0,0,0,1$. So if you multiply C^T into v this is going to be a this is what this is r cross 1 vector, this is a r cross 1 vector, your transverse is just a dot product you get a one cross one scalar in the end, right this is what you want to optimize or v_{2000} .

Subject to linear constraints of $sv=0$. You can have different objectives functions like minimize nutrient uptake, maximize metabolite production or efficient metabolite channeling and so on. We will look at these a little later. But you basically minimize or maximize some linear combination of the fluxes subject to your Stoichiometric mass balance constraints, right. And you can also do perturbation analysis you can do in silico, gene deletion, drug innovation so on and so forth. We will see what these are in a later class.

(Refer Slide Time: 18:48)

Simple illustration of FBA/LP



- ▶ Constraints define solution space
- ▶ Still a large number of solutions are possible
- ▶ Optimise for a desired goal: maximise NADH or ATP
- ▶ Obvious in this case: maximisation of NADH switches off ATP production, and vice-versa



So let us look at a very simple example, okay. So here you have reaction coming in like this A giving B+ATP A giving you know NADH+B on this direction so it is not exactly balanced in that sense but you can imagine that $r_A = r_B = x_1 + x_2$. This is again very similar to something that you studied in school. This kind of a circuit diagram which splits across Kirchhoff's Law, Kirchhoff's Law is basically charge balanced here we have mass balances, essentially a very same thing.

So here you have constraints I say $x_1 > 0, x_2 > 0$ and $x_1 + x_2 = r_A$. So because of that all your solutions lie on this line. Now you can say maximize x_1 which means will get to this point $x_1, 0$ or you know maximize, you know this point you will get to again $r_A, 0, r_A$ as the other point, right.

So here you basically have switching of one production and so on, right. So in this case the maximization of NADH switches off ATP production and maximization of ATP switches of NADH production, right. But you can have whatever appropriate objective function you want. So how do you choose the objective function is very interesting, I think we will stop here for this lecture and we will continue later.

(Refer Slide Time: 20:32)

Recap

Topics covered

- ▶ Steady-state Mass Balance
- ▶ Constraints
- ▶ Optimisation (LP)

In the next video ...

- ▶ The Objective Function
- ▶ Illustration of FBA

So in this video I hope you got an introduction to flux balance analysis particularly how you setup the steady state mass balance problem and how you add constraints to it to further reduce the solution space and finally setup an optimization problem to pick out one unique solution. In the next video, we will discuss FBA in little more detail.

And we will look at an objective function, how we setup the objective function for FBA and a very simple example to illustrate the various steps involved in FBA.