BIOMENG 261

TISSUE AND BIOMOLECULAR ENGINEERING

Module I: Reaction kinetics and systems biology

Oliver Maclaren
oliver.maclaren@auckland.ac.nz

LECTURE 9: FLUX BALANCE ANALYSIS CONTINUED

- Flux balance/constraint-based analysis continuued
- Null spaces and spans (linear algebra)
- Geometry of constraints
- Extra constraints
- Optimality conditions (linear programming)

MODULE OVERVIEW

Reaction kinetics and systems biology (*Oliver Maclaren*)
[12 lectures/4 tutorials/2 labs]

1. *Basic principles: modelling with reaction kinetics* [6 lectures]
   Physical principles: conservation, directional and constitutive. Reaction modelling. Mass action. Enzyme kinetics. Enzyme regulation. Mathematical/graphical tools for analysis and fitting.
2. *Systems biology I: overview, signalling and metabolic systems* [3 lectures]
   Overview of systems biology. Modelling signalling systems using reaction kinetics. Introduction to parameter estimation. Modelling metabolic systems using reaction kinetics. Flux balance analysis and constraint-based methods.
3. *Systems biology II: genetic systems* [3 lectures]
   Modelling genes and gene regulation using reaction kinetics. Gene regulatory networks, transcriptomics and analysis of microarray data.

RECALL: FLUX BALANCE ANALYSIS

Instead of the dynamic (ODE) problem, we aim to solve the *steady-state* equation

$$\mathbb{S}\mathbf{J} = \mathbf{0}$$

for the vector of *fluxes* $\mathbf{J}$, *here treated as unknown*.

- No consitutive equations/no rate parameters involved here.
- We don't need to know the metabolite concentrations, just solve for fluxes

RECALL: FLUX BALANCE ANALYSIS

For a given metabolic network there are *typically* (not always) more reactions than species/metabolites i.e.

More columns (unknowns) than rows (equations)

The problem is *underdetermined*, i.e. there are typically *multiple solutions*.

There is a non-trivial *null space*.

SPAN?

The *span* of a set of vectors is just the set of all linear combinations of these, i.e. the *hyperplane* these define.

Here we have

$$N(\mathbb{S}) = \text{span}\{\text{ indep. solutions of } \mathbb{S}\mathbf{J} = \mathbf{0}\}$$

Example.

NULL SPACE?

For a matrix $\mathbb{A}$ the *null space* is just the set of solutions to the zero problem

$$\mathbb{A}\mathbf{x} = \mathbf{0}$$

i.e. here

$$N(\mathbb{S}) = \{\mathbf{J} \mid \mathbb{S}\mathbf{J} = \mathbf{0}\}$$

Example.

UNIQUENESS? CONSTRAINT-BASED ANALYSIS

Clearly, there are multiple compatible solutions. To explore these further we can

- Add *bounds* (capacity constraints) on fluxes
- Add *directional* constraints (from thermodynamics)
- Look for special *'optimal'* solutions (e.g. maximum ATP production)

We say we are carrying out a *constraint-based analysis*...for obvious reasons! (FBA is a particular type of constraint-based analysis).

EXAMPLE

See handout.

GENERAL OPTIMISATION FRAMEWORK

We can formulate our problem as

$$\min z = \mathbf{c}^T\mathbf{J}$$

subject to

$$\mathbb{S}\mathbf{J} = \mathbf{0}$$

$$l_i \leq J_i \leq u_i$$

for $i = 1, ...N$ and a vector $\mathbf{c}$ of scalar 'costs' (weights), i.e. a…

GENERAL OPTIMISATION FRAMEWORK

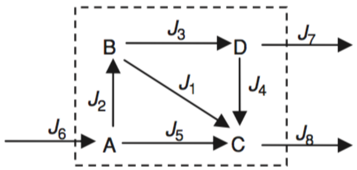
linear programming optimisation problem (see EngSci OpsRes courses!)

WHAT DO I NEED TO BE ABLE TO DO?

- Given a network, find $\mathbb{S}$
- Given an $\mathbb{S}$, draw a network
- Find the nullspace for a simple $\mathbb{S}$ (see handout)
- Describe/list some constraints or conditions that we might add to explore our null space and find special solutions
- Write down an optimisation problem given a problem description
- Solve a simple optimisation problem

EXTRA: BOUNDARIES, INTERNAL FLUXES AND INEQUALITY VS EQUALITY CONSTRAINTS

We often want to ‘draw boundaries’ around a ‘system’ of interest. We can either *include* these boundary fluxes as usual or treat them like *‘slack’* variables for *inequality* constraints



EXTRA: BOUNDARIES, INTERNAL FLUXES AND INEQUALITY VS EQUALITY CONSTRAINTS

Implicit inequality constraints give the polyhedra seen in:

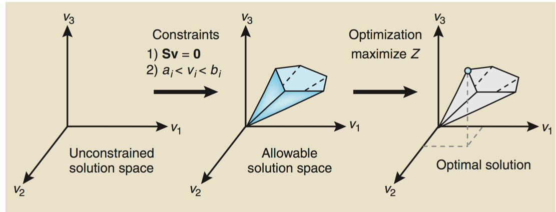


Figure 1 The conceptual basis of constraint-based modeling. With no constraints, the flux distribution of a biological network may lie at any point in a solution space. When mass balance constraints imposed by the stoichiometric matrix **S** (labeled 1) and capacity constraints imposed by the lower and upper bounds (a_i and b_i) (labeled 2) are applied to a network, it defines an allowable solution space. The network may acquire any flux distribution within this space, but points outside this space are denied by the constraints. Through optimization of an objective function, FBA can identify a single optimal flux distribution that lies on the edge of the allowable solution space.

EXTRA: BOUNDARIES, INTERNAL FLUXES AND INEQUALITY VS EQUALITY CONSTRAINTS

- *Equality* constraints $\mathbb{S}\mathbf{J} = \mathbf{0}$ define *hyperplanes* in the space of *all fluxes* (including boundary etc fluxes).
- *Inequality* constraints $\mathbb{S}\mathbf{J} \geq \mathbf{0}$ define *polyhedra* in the *reduced* set of fluxes (e.g. internal only).

Equivalent, given proper care, but just be aware of which.

Biomeng 261 Lecture 09

- Flux Balance Analysis (FBA) / Constraint-based analysis cont'd.

Some background math

- Null spaces & spans
- Geometry of constraints

$\Big\}$ Linear algebra

- Optimality conditions & optimisation problems

$\Big\}$ Linear programming

Recall: Determine $\underline{S}$ for the system

$$? \xrightarrow[\text{ATP} \to \text{ADP}]{J_1} \text{Gluc-6-P} \xrightarrow[\text{ATP} \to \text{ADP}]{J_2} ?$$

$$\text{Gluc-6-P} \xrightarrow{J_3} \text{Fruc-6-P}$$

$$\text{Fruc-6-P} \xrightarrow[\text{ATP} \to \text{ADP}]{J_4} \text{Fruc-1,6-P}_2 \xrightarrow{J_5} ?$$

$$\text{ADP} \xrightarrow{J_6} \text{ATP}$$

$$\text{ATP} \xrightarrow{J_7} \text{ADP}$$

$$\text{ATP} + \text{AMP} \xrightarrow{J_8} 2\,\text{ADP}$$

Answer:

$$\underline{S} = \begin{array}{c} \\ \\ \\ \\ \\ \\ \\ \end{array}\begin{array}{cccccccc} J_1 & J_2 & J_3 & J_4 & J_5 & J_6 & J_7 & J_8 \end{array}$$

$$\underline{S} = \begin{pmatrix} 1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 \\ -1 & -1 & 0 & -1 & 0 & 1 & -1 & -1 \\ 1 & 1 & 0 & 1 & 0 & -1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix} \begin{array}{l} \leftarrow \text{Gluc-6-P} \\ \leftarrow \text{Fruc-6P} \\ \leftarrow \text{Fruc-1,6-P}_2 \\ \leftarrow \text{ATP} \\ \leftarrow \text{ADP} \\ \leftarrow \text{AMP} \end{array}$$

Recall

What's the catch? $\underline{S}\bar{J} = \bar{0}$ soln?

Here:

6 rows $\longleftarrow$ equations/constraints

8 columns $\longleftarrow$ unknowns (J)

$\boxed{\text{unknowns} > \text{eq}^{\text{n}}\text{s}}$

In general

- more reactions/fluxes (unknowns) than metabolites/concentrations (equations)
- often don't know all metabolites involved (lose a rows)

Shape:

$$\begin{array}{c} \\ C_1 \\ \vdots \\ C_M \end{array} \begin{array}{c} J_1 \; \cdots \cdots \; J_N \\ \left[\begin{array}{c} \\ \longleftarrow \text{wider} \longrightarrow \\ \\ \end{array} \right] \end{array}$$

$N > M$

$\rightarrow$ $\underline{S}\bar{J} = \bar{0}$ is usually underdetermined

$\rightarrow$ ie multiple solutions

$\rightarrow$ makes sense since only using conservation of mass.

Recall

Null spaces (of $\underline{S}$ say)

- The nullspace of a matrix $\underline{A}$ is the set of all solutions to $\underline{A}\bar{x} = \bar{0}$

FBA: $\underline{A} = \underline{S}$, $\bar{x} = \bar{J}$

- zero vector is always in nullspace
  $\underline{A}\bar{0} = \bar{0}$
- A non-trivial null-space is when we have non-zero solutions in the null space
- dimension of null-space is # vars − # indep. constraints

Mathematically:

$$\boxed{\underbrace{N(\underline{A})}_{\text{null space of } \underline{A}} = \underbrace{\{}_{\text{set of}} \overbrace{\bar{x}}^{\text{vectors}} \underbrace{|}_{\text{such that}} \overbrace{\underline{A}\bar{x} = \bar{0}}^{\text{they satisfy condition.}} \}}$$

---

recall

Example

$\underline{S}\,\bar{J} = \bar{0}$: $\begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} J_1 \\ J_2 \\ J_3 \\ J_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

4 unknowns $(J_1, .. J_4)$

2 eq^{n}s

expect $4-2 = 2$ free vars

$\Rightarrow$ two indep. sol^{n}s.

① $J_1 - J_2 = 0$

② $-J_2 + J_3 - J_4 = 0$

choose J_1 free

① $\rightarrow$ gives J_2

② & above : choose one of J_3 or J_4

$\rightarrow J_3 \rightarrow$ determines J_4

so J_1, J_3 free $\rightarrow J_2, J_4$ dependent.

& $J_2 = J_1$

$\underbrace{J_4}_{J_2, J_4} = \underbrace{-J_2 + J_3}_{} = \underbrace{-J_1 + J_3}_{J_1, J_3}$

recall

Independent vector sol^{n}s ?

$\Rightarrow$ All sol^{n}s have form $\begin{pmatrix} J_1 \\ J_1 \\ J_3 \\ J_3 - J_1 \end{pmatrix}$

Can generate two independent vectors

by setting $\left\{ \begin{matrix} J_1 = 1, \; J_3 = 0 \\ J_1 = 0, \; J_3 = 1 \end{matrix} \right\}$ in turn

ie

$\bar{J}^{(1)} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ -1 \end{pmatrix}$

& $\bar{J}^{(2)} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$

procedure ensures neither can be written as linear combo of other(s)

$\rightarrow$ all other sol^{n}s can be written as $\bar{J} = a\bar{J}^{(1)} + b\bar{J}^{(2)}$ for some a, b.

$\{\bar{J}^{(1)}, \bar{J}^{(2)}\}$ form 'basis' for null space.

Leads to idea of ...

Span : $\boxed{\text{span}\{\bar{J}^{(1)}, \bar{J}^{(2)}, \ldots \bar{J}^{(n)}\}}$ $\boxed{a\bar{J}^{(1)} + b\bar{J}^{(2)} + \ldots}$

- Set of all linear combinations of a set of vectors
- Every vector in the null space can be written as a linear combination of independent solutions to $\underline{S}\bar{J} = \bar{0}$

ie

$$\boxed{N(\underline{S}) = \text{span}\left\{ \begin{array}{c}\text{independent vectors} \\ \text{solving } \underline{S}\bar{J} = \bar{0}\end{array} \right\}}$$

where $\boxed{\text{span}\{\bar{J}^{(1)}, \bar{J}^{(2)}, \ldots\}}$

is short for

$$\boxed{\{ a\bar{J}^{(1)} + b\bar{J}^{(2)} + \ldots \mid \text{for all } a, b, \ldots \}}$$

Procedure :

1. – reduce to minimal set of eq^{n}s
2. – choose free vars (eg $m-n$ of them)
3. – find implied set of independent vectors
4. – write $N(\underline{S}) = \text{span}\{ ? \}$

Two free vars $\Rightarrow$ Two indep. vectors etc.

Huh? Best illustrated via (more) examples.

Another Example

$n = 6$

$$m=4 \quad \begin{bmatrix} -1 & 1 & 0 & 0 & 0 & 0 \\ 2 & -1 & -2 & 0 & 0 & 0 \\ 0 & 0 & 3 & -3 & 3 & 0 \\ 0 & 0 & 0 & 1 & -1 & -1 \end{bmatrix}$$

Expect $n - m = 6 - 4 = 2$ free vars &
hence 2 independent vectors
(if all eq^{n}s independent → I'll usually make this true!)

Approach: use elementary row ops to make upper triangular...

OR

this course usually! { just expand out & solve for in terms of free vars by being sensible

---

Then ↳ [setting each free var non-zero & rest as zero in turn generates independent vector solutions.]

Example cont'd

$$\begin{bmatrix} -1 & 1 & 0 & 0 & 0 & 0 \\ 2 & -1 & -2 & 0 & 0 & 0 \\ 0 & 0 & 3 & -3 & 3 & 0 \\ 0 & 0 & 0 & 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} J_1 \\ J_2 \\ J_3 \\ J_4 \\ J_5 \\ J_6 \end{bmatrix} = \bar{0}$$

$$\Rightarrow \quad -J_1 + J_2 = 0 \quad (1)$$

$$2J_2 - J_2 - 2J_3 = 0 \quad (2)$$

$$\cancel{3}J_3 - \cancel{3}J_4 + \cancel{3}J_5 = 0 \quad (3)$$

$$J_4 - J_5 - J_6 = 0 \quad (4)$$

---

4 equations, 6 vars.
choose 2 & make sure rest are determined

---

eg choose J_1 $\xrightarrow{(1)}$ gives J_2
$\xrightarrow{(2)}$ gives J_3

Leaves eg J_4 or J_5 ← choose

Choosing J_1 & J_5 as free gives

$$J_2 = J_1$$
$$J_3 = \tfrac{1}{2} J_2 = \tfrac{1}{2} J_1$$
$$J_4 = J_3 + J_5 = \tfrac{1}{2} J_1 + J_5$$
$$J_6 = J_4 - J_5 = \tfrac{1}{2} J_1$$

Make J_1 & J_5 non-zero in turn to get independent sol^{n}s eg

① $J_1 = 2,\ J_5 = 0$

sol^n 1. $\bar{J}^{(1)} = \begin{pmatrix} 2 \\ 2 \\ 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}$

② $J_1 = 0,\ J_5 = 1$

sol^n 2 $\bar{J}^{(2)} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}$

Finally, write

$$\boxed{N(\underset{\sim}{S}) = \text{span}\{\bar{J}^{(1)}, \bar{J}^{(2)}\}}$$

where $\bar{J}$s are as above

(phew!)

Summary of example:

$$N(\underset{\sim}{S}) = \text{span}\left\{ \begin{pmatrix} 2 \\ 2 \\ 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right\}$$

→ a 2D 'subspace' of a 6D flux space.

→ all possible sol^{n}s 'live' here.

via:

6 variables, 4 equations

⇒ 6-4 = 2 free variables (entries in vectors)

⇒ 2 independent vectors in span

(eg set each free entry to 1 & others to 0 in turn → generates same number of vectors as free variables).

Special solutions - uniqueness?

we have linear constraints on fluxes

(would usually be nonlinear in concentrations eg using mass action)

→ These define a nontrivial nullspace, ie a hyperplane ($\underset{\sim}{S}\,\bar{J} = \bar{0}$ form)

(or a polyhedral feasible region if written $\underset{\sim}{S}\,\bar{J} \geq \bar{0}$)

↳ see appendix & 'internal' fluxes

→ we might want to look at ways to 'pick out' particular solutions within this space

→ add $\left\{ \begin{array}{l} \text{extra constraints} \\ \text{objectives to max/min.} \end{array} \right.$

Constraints & optimality conditions

we can add many types of extra constraints or conditions to further narrow down our space of possible sol^{n}s

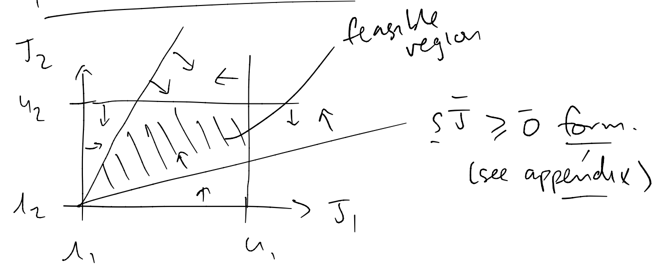
- bounds/signs of fluxes
- thermodynamic feasibility (directional constraints)
- optimality or extreme cases (eg maximise energy production)

o Most natural next constraints: lower & upper flux bounds

$\boxed{\text{Capacity constraints}}$ (bounds)

$$\boxed{l_i \leq J_i \leq u_i}$$

Gives eg



o For irreversible, use $\boxed{J_i \geq 0}$ (why?)

(thermo)

'Optimal' solutions

A useful way to pick out 'special' solutions is to use optimality (max/min) conditions

→ tells us about limits on what is possible

eg "maximal rate of ATP production is --- "

Note: a real system may or may not reach these limits!

- many competing goals, so may not be optimal for any one goal
- still useful way to predict & understand
  - ↳ make prediction then test experimentally!
- many sensible constraints can be re-written as max/min conditions

Example: cont'd

Recall previous example in terms of J_1 & J_5 as free vars

$$\boxed{\begin{aligned} J_2 &= J_1 \\ J_3 &= J_1/2 \\ J_4 &= J_1/2 + J_5 \\ J_6 &= J_1/2 \end{aligned}}$$

ie $J_2, J_3, J_4, J_6 = f(J_1, J_5)$

6 vars, 4 eq^n, 2 free, 4 det.

just $S\bar{J} = \bar{0}$ rewritten

To find a particular sol^n

- suppose $2 \leq J_1, J_5 \leq 10$ } bounds apply to both
- $J_i \geq 0$ for all $i = 1, \dots 6$
- J_4 is ATP production
- J_6 is lactate production } for example

Goals: Case 1: Max ATP production

Case 2: Max ATP production while minimising lactate prod.

Key idea: re-write objective in terms of free vars as well

Case I max J_4

subject to

- $\underline{S}\,\bar{J} = \bar{0}$
- $2 \leq J_1 \leq 10$
- $2 \leq J_5 \leq 10$

$\rightarrow$ we rewrote as

$J_2 = J_1$
$J_3 = J_1/2$
$J_4 = J_1/2 + J_5$
$J_6 = J_1/2$

Sol^n use $J_4 = J_1/2 + J_5$ } obj. function in terms of free vars

want to maximise the above

$\rightarrow$ make both as big as possible } free to vary up to bounds

$\Rightarrow J_1 = 10,\ J_5 = 10$

$\Rightarrow J_4 = 10/2 + 10 = 15$

$\Rightarrow J_2 = 10,\ J_3 = 5,\ J_6 = 5$

$$\boxed{\text{So} \quad \bar{J}^{sol} = \begin{pmatrix} 10 \\ 10 \\ 5 \\ 15 \\ 10 \\ 5 \end{pmatrix}}$$

Case 2. [Note: $\min J \equiv \max -J$]

$$\Rightarrow \boxed{\max\ aJ_4 - bJ_6}$$

for some $a, b > 0$ weights
(relative 'value' of each)

subject to
same constraints as before.

Sol^n: J_1 & J_5 free

& $J_4 = J_1/2 + J_5$, $J_6 = J_1/2$

So $\underbrace{aJ_4 - bJ_6 = \left(\frac{a-b}{2}\right).J_1 + aJ_5}_{\text{want to maximise}}$ } obj. in terms of free variables.

If $a > b$ then as before, $J_1 = J_5 = 10$ etc.

If $a < b$ then $J_1 \rightarrow$ lower bound instead } why?

$J_1 = 2,\ J_5 = 10$ (& solve for $J_2, .. J_6$)

If $a = b$ $\underbrace{J_1 = ?}_{\text{still free}}$ $J_5 = 10$ } not enough info for unique sol^n

General optimisation framework

$\boxed{\min}\ z = \bar{c}^T \cdot \bar{J} = (c_1\ c_2\ \dots\ c_n) \begin{pmatrix} J_1 \\ J_2 \\ \vdots \\ J_n \end{pmatrix}$

$= c_1 J_1 + c_2 J_2 + \dots$ } objective function

$\boxed{\text{subject to}}$

$\underset{\sim}{S}\bar{J} = 0$

$l_i \leq J_i \leq u_i$ } constraints

---

z : scalar (number), total 'cost' (since min)

$\bar{c}$: vector of weights (here 'costs' since min)

eg $z = \bar{c}^T\bar{J} = (1\ \ 2\ \ -1)\begin{pmatrix} J_1 \\ J_2 \\ J_3 \end{pmatrix}$

$= J_1 + 2J_2 - J_3$

{ Linear objective function + Linear constraints } "Linear Programming" └ EngSci 255 etc.

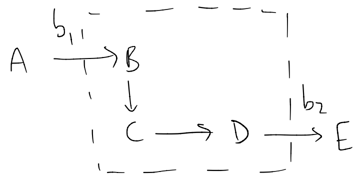
Um? What should I be able to do?

- Given network, find $\underset{\sim}{S}$
- Given $\underset{\sim}{S}$, draw network
- Find nullspace for simple $\underset{\sim}{S}$
  - └ reduce to independent eq^{n}s } usually done already
  - └ choose free fluxes
  - └ find implied independent vectors
  - └ write as span { ↓ }
- Describe typical/useful additional constraints } bounds, directions
- Write down optimisation problem for given description } max/min
- Solve very simple optimisation problems (by writing in terms of free vars etc as shown)

Appendix: Boundary vs Internal Fluxes

- Sometimes we want to draw 'system boundaries' & call some fluxes 'boundary' fluxes & some 'internal' fluxes

eg



b: boundary
J: internal
(fluxes).

We can either include these as usual:

$$\begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{cccc} J_1 & J_2 & b_1 & b_2 \end{array} \\ \underbrace{\begin{bmatrix} 0 & 0 & -1 & 0 \\ -1 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}} \end{array}$$

gives $\boxed{\underline{S}\bar{J} = \bar{0}}$

OR

focus on internal:

$$\begin{array}{c} \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cc} J_1 & J_2 \end{array} \\ \underbrace{\begin{bmatrix} -1 & 0 \\ 1 & -1 \\ 0 & 1 \end{bmatrix}} \end{array}$$

gives $\boxed{\underline{S}\bar{J} \geq 0}$ if we choose signs carefully

$(\underline{S}\bar{J} = \bar{b}_{net} \geq 0)$

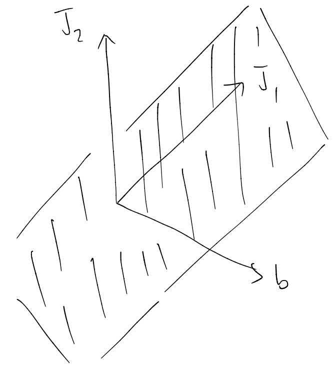
Appendix: Geometry of Nullspace vs Feasible region

(Nullspace: $=$; Feasible region: $\geq$)

The two forms

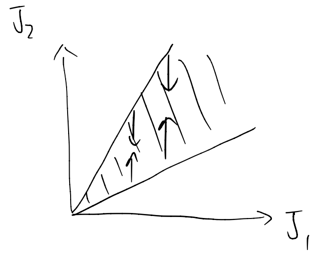
$\underbrace{\underline{S}\bar{J} = \bar{0}}_{\text{all fluxes}}$ or $\underbrace{\underline{S}\bar{J} \geq \bar{0}}_{\text{subset of fluxes, eg internal only}}$

are equivalent but have slightly different geometric pictures



$\underline{S}\bar{J} = \bar{0}$

eg 2D plane in 3D



$\underline{S}\bar{J} \geq 0$

eg 2D cone

b becomes 'slack' var.

eg 3−1 = 2D plane

∘ Nullspace

2 vars, 2 inequality constraints

2D cone/polyhedra

∘ Feasible region

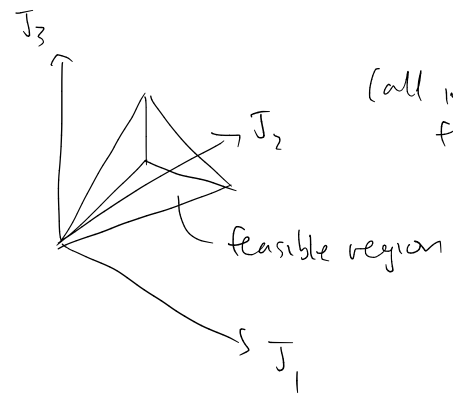
Appendix: what/why?

In general we will use 'standard form'

$$\boxed{\underline{S}\,\bar{J} = \bar{0}}$$

& use $\boxed{\text{nullspace}}$ (ie include boundary fluxes)

But you may see pictures like



(all internal fluxes)

which come from inequality constraint version → just be aware

---

(Aside: a general idea:
simple in higher dim or complex in lower dim)
