$\boxed{6}$ Finite dynamical systems & Computational algebra prelims

<u>Reverse engineering</u>: Start with experimental data and build a model. "Discover" the network through experiments & observations.

Typically, there is insufficient data to uniquely infer a gene regulatory network (GRN) model.

<u>Model selection</u> is needed - select biologically feasible or likely ones.

<u>Example</u>: Input states $\vec{s}_1, \dots, \vec{s}_m \in \mathbb{F}^n$ where $\mathbb{F} = \{0,1\}$

Output states $\vec{t}_1, \dots, \vec{t}_m \in \mathbb{F}^n$

That is, $F(\vec{s}_i) = (f_1(\vec{s}_i), f_2(\vec{s}_i), \dots, f_n(\vec{s}_i))$

$$\begin{array}{ccccc}
\vec{s}_1 = (s_{11}, s_{12}, \dots s_{1n}) & \vec{s}_2 = (s_{21}, \dots, s_{2n}) & \dots & \vec{s}_m = (s_{m1}, \dots, s_{mn}) \\
\downarrow F & \downarrow F & & \downarrow \\
\vec{t}_1 = (t_{11}, t_{12}, \dots, t_{1n}) & \vec{t}_2 = (t_{21}, \dots, t_{2n}) & \dots & \vec{t}_m = (t_{m1}, \dots, t_{mn})
\end{array}$$

This is just a few transitions in the phase space (size 2^n).

<u>Goal</u>: Recover the actual functions $F = (f_1, \dots, f_n)$ and the wiring diagram.

Framework:

Def: A finite dynamical system (FDS) is a function $F = (f_1, \dots, f_n): S^n \to S^n$ where each $f_i: S^n \to S$ is a local function, and $|S| < \infty$.

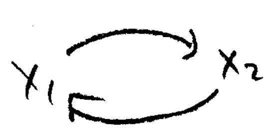
Theorem: If $S = \mathbb{F}$, a finite field (e.g., $\mathbb{Z}_2, \mathbb{Z}_3, \mathbb{Z}_p$, etc.), then any function $f_i: \mathbb{F}^n \to \mathbb{F}$ is a polynomial in $x_1, \dots, x_n$.

If $S = \mathbb{F}$, we call an FDS a polynomial dynamical system (PDS).

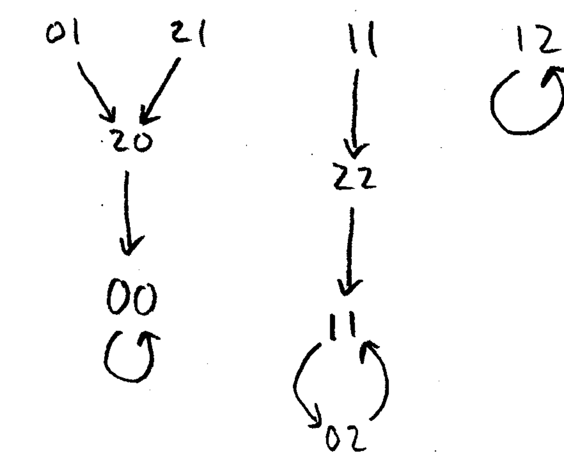
Ex: Let $\mathbb{F} = \mathbb{Z}_3 = \{0,1,2\}$ and $F = (f_1, f_2): \mathbb{F}^2 \to \mathbb{F}$ be a PDS with local functions

$$f_1 = f_1(x_1, x_2) = 2x_2$$
$$f_2 = f_2(x_1, x_2) = x_1 + x_1^2$$

wiring diagram



phase space (size $|\mathbb{F}|^2 = 9$)



We'll need multivariate polynomial division

Recall: (single variable) polynomial division is well-defined.

$$a(x) \;=\; q(x)\;\; b(x) \;+\; r(x)$$

$$x^4-3x^3-x^2+2 = (x^2-5x+10)(x^2+2x-1) + (-25x-12)$$

$\uparrow$ quotient $\qquad\qquad$ $\uparrow$ remainder

$$\begin{array}{r} x^2-5x+10 \\ x^2+2x-1\,\overline{\big)\,x^4-3x^3-x^2+0x+2} \\ \underline{x^4+2x^3-x^2\qquad\quad} \\ -5x^3+0x^2+0x \quad \\ \underline{-5x^3-10x^2+5x\quad} \\ 10x^2-5x+2 \\ \underline{10x^2+20x-10} \\ \text{"remainder"}\quad -25x-12 \end{array}$$

This works because there is a total ordering on monomials:

$$\cdots > x^{m+1} > x^m > \cdots > x^2 > x > 1$$

Multivariate polynomial division isn't well-defined (yet).

Example: $f = x^2$, $g = x^2 + xy$. Divide f by g:

$$\begin{array}{r} 1 \\ x^2+xy^2\,\overline{\big)\,x^2+0xy^2} \\ \underline{x^2+\;xy^2} \\ -xy^2 \end{array} \qquad \text{vs.} \qquad \begin{array}{r} 0 \\ xy^2+x^2\,\overline{\big)\,0xy^2+x^2} \\ \underline{0xy^2+0x^2} \\ x^2 \end{array}$$

So $\quad x^2 = 1\,(x^2+xy^2) + (-xy^2) \quad$ vs. $\quad x^2 = 0\,(xy^2+x^2) + x^2$

$$f(x,y) = 1\cdot g(x,y) + (-xy^2) \quad \text{vs.} \quad f(x,y) = 0\cdot g(x,y) + (-xy^2).$$

We need to define an ordering on the set of <u>monomials</u>.

(e.g., which comes "first": x^2yz, xy^3z, or xyz^2?)

Let K be a field (e.g., $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, $\mathbb{Z}_p$).

$K[x_1,\dots,x_n]$ is a <u>polynomial ring</u> (set of all multivariate polynomials in variables $x_1,\dots,x_n$.)

A <u>polynomial</u> $f(x_1,\dots,x_n)$ is a sum of <u>monomials</u> $X^\alpha = x_1^{\alpha_1}x_2^{\alpha_2}\cdots x_n^{\alpha_n}$, where $\alpha = (\alpha_1,\alpha_2,\dots,\alpha_n)$.

<u>Need</u>: A <u>total ordering</u> on monomials, i.e., for every pair X^α & X^β, exactly one of these must be true:

$$X^\alpha < X^\beta, \quad X^\alpha = X^\beta, \quad X^\alpha > X^\beta.$$

<u>Def</u>: A <u>monomial ordering</u> on $K[x_1,\dots,x_n]$ is any relation $>$ on $\mathbb{Z}_{\geq 0}^n := \{(a_1,\dots,a_n) : a_i \in \mathbb{Z}_{\geq 0}\}$ satisfying:

1. $>$ is a total ordering on $\mathbb{Z}_{\geq 0}^n$
2. If $\alpha > \beta$ and $\gamma \in \mathbb{Z}_{\geq 0}^n$, then $\alpha + \gamma > \beta + \gamma$.
3. $>$ is a well-ordering on $\mathbb{Z}_{\geq 0}^n$, i.e., every nonempty subset of $\mathbb{Z}_{\geq 0}^n$ has a smallest element.

Examples: Let $\alpha = (\alpha_1, \dots, \alpha_n)$, $\beta = (\beta_1, \dots, \beta_n) \in \mathbb{Z}^n_{\geq 0}$.

(1) lexicographic (lex) "dictionary order" $\alpha >_{lex} \beta$ if the left-most non-zero entry of $\alpha - \beta$ is positive.

(2) graded lexicographic (grlex) "total degree first; breaks ties with lex"

$\alpha >_{grlex} \beta$ if $|\alpha| > |\beta|$ (that is, $\sum_{i=1}^{n} \alpha_i > \sum_{i=1}^{n} \beta_i$), or $|\alpha| = |\beta|$ and $\alpha >_{lex} \beta$.

(3) graded reverse lexicographic (grevlex): $\alpha >_{grevlex} \beta$ if $|\alpha| > |\beta|$, or $|\alpha| = |\beta|$ and the right-most non-zero entry of $\alpha - \beta$ is negative.

Examples:

- For all 3 of these orders,

$$\begin{array}{ccccccc} (1,0,\dots,0) & > & (0,1,0,\dots,0) & > \cdots > & (0,\dots,0,1) \\ x_1 & > & x_2 & > \cdots > & x_n \end{array}$$

- $x^2 >_{lex} xy^2 >_{lex} y^2$, since $(2,0) >_{lex} (1,3) >_{lex} > (0,2)$.

- $y^2 z <_{grlex} xz^2$ but $y^2 z >_{grevlex} xz^2$ $[(0,2,1)$ vs. $(1,0,2)]$.

- $x^5yz >_{grlex} x^4yz^2$ and $x^5yz >_{grevlex} x^4yz^2$ $[(5,1,1)$ vs. $(4,1,2)]$.

Even with a monomial order, there are problems with multivariate polynomial division.

<u>Ideal membership problem</u>: Given f, and $f_1, \dots, f_m$, are there polynomials $h_1, \dots, h_m$ such that $f = h_1 f_1 + \dots + h_m f_m$?

Compare to "subspace membership problem" Given $v \in V$ and basis $v_1, \dots, v_m$, are there constants $a_1, \dots, a_m \in \mathbb{R}$ s.t. $v = a_1 v_1 + \dots + a_m v_m$?

<u>Def</u>: Let K be a field. The set

$$I = \langle f_1, \dots, f_m \rangle = \{ h_1 f_1 + \dots + h_m f_m : h_i \in K[x_1, \dots, x_n] \}$$

is called an <u>ideal</u> in $K[x_1, \dots, x_n]$, <u>generated</u> by $f_1, \dots, f_m$.

<u>Question</u>, reformulated: Is f in I?

<u>Annoyances</u>

- The order in which f is divided by the f_i's affects the remainder.
- A nonzero remainder does not imply $f \notin I$.
- The generating set is not unique, nor is its size. (Unlike subspaces in linear algebra!)

Examples (of annoyances):

- Order matters: Let $f = x^2y + xy^2 + y^2$, $f_1 = xy - 1$, $f_2 = y^2 - 1$

Use lex with $x > y$, divide f_1 into f, then f_2:

Result: $$\underbrace{x^2y + xy^2 + y^2}_{f} = \underbrace{(x+y)}_{h_1} \cdot \underbrace{(xy-1)}_{f_1} + \underbrace{1}_{h_2} \cdot \underbrace{(y^2-1)}_{f_2} + \underbrace{x+y+1}_{r}$$

Instead, if we divide f_2 into f and then f_1:

Result: $$\underbrace{x^2y + xy^2 + y^2}_{f} = \underbrace{(x+1)}_{h_1} \cdot \underbrace{(y^2-1)}_{f_1} + \underbrace{x}_{h_2} \underbrace{(xy-1)}_{f_2} + \underbrace{2x+1}_{r}$$

- nonzero remainders: Let $f = xy^2 - x$, $f_1 = xy + 1$; $f_2 = y^2 - 1$.

Dividing f by f_1 then f_2 yields

$$\underbrace{xy^2 - x}_{f} = \underbrace{y}_{h_1} \cdot \underbrace{(xy+1)}_{f_1} + \underbrace{0}_{h_2} \cdot \underbrace{(y^2-1)}_{f_2} + \underbrace{(-x-y)}_{r}$$

Dividing f by f_2 then f_1 yields

$$\underbrace{xy^2 - x}_{f} = \underbrace{x}_{h_1} \cdot \underbrace{(y^2-1)}_{f_1} + \underbrace{0}_{h_2} \cdot \underbrace{(xy+1)}_{f_2} + 0 \quad \Rightarrow \quad f \in \langle f_1, f_2 \rangle.$$

- <u>Size</u> of minimal generating sets.

In $\mathbb{R}[x,y]$, the ideal $\langle x,y\rangle = \langle x+xy,\ y+xy,\ x^2,\ y^2\rangle$ (Exercise)

Fortunately, it is possible to choose a "special basis" that avoids these annoyances.

<u>Theorem</u> An ideal I in $K[x_1,\dots,x_n]$ has a special generating set $\{g_1,\dots,g_t\}$ such that the remainder of polynomial division of f by $f_1,\dots,f_m$ (in any order) is zero iff $f \in \langle f_1,\dots,f_m\rangle$

Such a generating set is called a <u>Gröbner basis</u>.

<u>Def</u>: Given a polynomial f and ordering $>$, the <u>leading monomial</u> of f, denoted $in_>(f)$, is the monomial ordered first by $>$ (assume monic).

<u>Def</u>: The <u>initial ideal</u> of I, is the set

$$in_>(I) := \langle in_>(f) : f \in I\rangle.$$

<u>Def</u>: A finite subset $\mathcal{G} \subseteq I$ is a <u>Gröbner basis</u> wrt $>$ if

$$\boxed{in_>(I) = \langle in_>(g) : g \in \mathcal{G}\rangle}$$

<u>Example</u>: Use grlex, let $x > y$, and $f_1 = x^3 - 2xy$
$f_2 = x^2y - 2y^2 + x$.

The ideal $I = \langle f_1, f_2 \rangle \subseteq K[x, y]$.

$in_>(f_1) = x^3$, $in_>(f_2) = x^2y$, so $\langle x^3, x^2y \rangle \subseteq in_>(I)$.

But $in_>(I)$ contains more!

$$\begin{aligned} x f_2 - y f_1 &= x(x^2y - 2x^2 + x) - y(x^3 - 2xy) \\ &= x^3y - 2xy^2 + x^2 - x^3y - 2xy^2 \\ &= x^2 \in I. \end{aligned}$$

Clearly, $in_>(x^2) = x^2 \implies x^2 \in in_>(I)$, but $x^2 \notin \langle x^3, x^2y \rangle$.

So $\{f_1, f_2\}$ is <u>not</u> a Gröbner basis for I!

Using Sage, we get that the following <u>is</u> a GB:

$$\{x^3 - 2xy,\ x^2y - 2y^2 + x,\ x^2,\ xy,\ y^2 - \tfrac{1}{2}x\}$$

So $$\begin{aligned} in_>(I) &= \langle x^3, x^2y, x^2, xy, y^2 \rangle \\ &= \langle x^2, xy, y^2 \rangle. \end{aligned}$$

The monomials in I that are <u>not</u> in $in_>(I)$ are called

<u>standard monomials</u>.