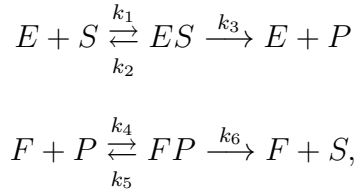


Clase 26/09/18: Métodos algebraicos en Singular para el estudio de redes bioquímicas.

Algunos ejemplos y Ejercicios

Consideremos el siguiente sistema de reacciones químicas bajo cinética de acción de masas:



donde k_1, \dots, k_6 son constantes (de reacción) *positivas*.

Notemos con corchetes la concentración de cada especie y establezcamos la siguiente correspondencia: $x_1 = [S]$, $x_2 = [P]$, $x_3 = [E]$, $x_4 = [F]$, $x_5 = [ES]$, $x_6 = [FP]$.

Las ecuaciones entonces son

$$\begin{aligned}\dot{x}_1 &= -k_1 x_1 x_3 + k_2 x_5 + k_6 x_6 \\ \dot{x}_2 &= -k_4 x_2 x_4 + k_5 x_6 + k_3 x_5 \\ \dot{x}_3 &= -k_1 x_1 x_3 + (k_2 + k_3) x_5 \\ \dot{x}_4 &= -k_4 x_2 x_4 + (k_5 + k_6) x_6 \\ \dot{x}_5 &= -\dot{x}_3 \\ \dot{x}_6 &= -\dot{x}_4.\end{aligned}$$

De estas dos últimas deducimos que existen constantes C_1 y C_2 tales que $x_3(t) + x_5(t) = C_1$ y $x_4(t) + x_6(t) = C_2$ para todo tiempo t . En particular, $C_1 = x_3(0) + x_5(0) \geq 0$ y $C_2 = x_4(0) + x_6(0) \geq 0$.

También vale que $\dot{x}_1 + \dot{x}_2 + \dot{x}_5 + \dot{x}_6 = 0$, por lo que existe una constante C_3 (≥ 0) tal que $x_1(t) + x_2(t) + x_5(t) + x_6(t) = C_3$ para todo tiempo t . A estas tres últimas ecuaciones que dedujimos se las conoce como *leyes de conservación*.

Estamos interesados, en particular, en calcular los *estados estacionarios* del sistema. Es decir, los $x^* = (x_1^*, \dots, x_6^*) \in \mathbb{R}_{\geq 0}^6$ que satisfagan $\dot{x}_1 = \dots = \dot{x}_6 = 0$. Y además nos interesan aquellos que cumplan las leyes de conservación para ciertos valores de C_1, C_2, C_3 .

Esto nos lleva a querer resolver el sistema de ecuaciones **polinomiales**

$$\begin{cases} \dot{x}_1 = 0 \\ \dot{x}_3 = 0 \\ \dot{x}_4 = 0 \\ x_3 + x_5 = C_1, x_4 + x_6 = C_2, x_1 + x_2 + x_5 + x_6 = C_3. \end{cases} \quad (1)$$

- No siempre se pueden hallar las soluciones a un sistema polinomial.

- Nos gustaría, además, no tener que asignarle valores específicos a $k_1, \dots, k_6, C_1, C_2$ y C_3 , para poder conseguir resultados más generales.
- ¿Podemos, al menos, saber si el sistema tiene una sola solución o más de una solución? Tal vez no tenga ninguna. Y si las tiene, ¿son todas biológicamente significativas?
- Aunque no las podamos hallar explícitamente, ¿podemos igual encontrar alguna característica particular de ellas?

Observación 1:

$$(k_2 + k_3)\dot{x}_1 - k_2\dot{x}_3 = -k_1(k_2 + k_3)x_1x_3 + k_6(k_2 + k_3)x_6 + k_1k_2x_1x_3 = -k_1k_3x_1x_3 + k_6(k_2 + k_3)x_6,$$

$$\underbrace{(k_5 + k_6)}_{>0} \underbrace{[(k_2 + k_3)\dot{x}_1 - k_2\dot{x}_3]}_{>0} - k_6(k_2 + k_3)\dot{x}_4 = -k_1k_3(k_5 + k_6)x_1x_3 + k_4k_6(k_2 + k_3)x_2x_4.$$

Obtenemos, por medio de esta operación lineal entre las ecuaciones, el sistema equivalente:

$$-k_1k_3(k_5 + k_6)x_1x_3 + k_4k_6(k_2 + k_3)x_2x_4 = 0, \tag{2}$$

$$-k_1x_1x_3 + (k_2 + k_3)x_5 = 0, \tag{3}$$

$$-k_4x_2x_4 + (k_5 + k_6)x_6 = 0, \tag{4}$$

$$x_3 + x_5 = C_1, x_4 + x_6 = C_2, x_1 + x_2 + x_5 + x_6 = C_3. \tag{5}$$

En este nuevo sistema hemos *eliminado*, de la primera ecuación, las variables x_5 y x_6 .

Observación 2: Si x^* es una solución tal que $x_3^* = 0$, de (3) deducimos que $x_5^* = 0$ (porque $k_2 + k_3 > 0$), y por (5) deducimos que $C_1 = 0$. Sin embargo, queremos considerar sistemas donde la cantidad total inicial de enzima sea positiva, es decir con $C_1 > 0$. Luego, sólo nos interesan aquellas soluciones x^* con $x_3^* \neq 0$.

Observación 3: De (2) podemos despejar

$$x_1 = \frac{k_4k_6(k_2 + k_3)}{k_1k_3(k_5 + k_6)} \cdot \frac{x_2x_4}{x_3},$$

y juntando esto con (3) tenemos

$$x_5 = \frac{k_4k_6}{k_3(k_5 + k_6)} x_2x_4.$$

Finalmente, de (4) vemos que

$$x_6 = \frac{k_4}{k_5 + k_6} x_2x_4.$$

Podemos luego reemplazar estos valores en (5) y el sistema se reduce a uno *no lineal* de 3 ecuaciones con 3 incógnitas (y 9 parámetros).

Observación 4: De lo anterior y (5) vemos que $x_4 \left(1 + \frac{k_4}{k_5 + k_6} x_2\right) = C_2$, por lo que

$$x_4 = \frac{C_2}{1 + \frac{k_4}{k_5+k_6}x_2} \underset{x_2 \leq C_3}{\geq} \frac{C_2}{1 + \frac{k_4}{k_5+k_6}C_3}.$$

En breve, haremos las cuentas con ayuda de Singular [3], lo que nos permitirá también analizar sistemas más complejos. Pero primero veamos una herramienta matemática que nos será de gran utilidad para poder eliminar variables.

Bases de Gröbner (un poco de teoría matemática):

Hablemos brevemente del concepto de base de Gröbner por medio de un ejemplo muy sencillo. Para un buen desarrollo de este tema ver [1, 2].

Si tenemos los monomios x_1x_2 y x_1x_3 , ¿cuál es mayor? Hay varios *órdenes monomiales* posibles. El más conocido es el *lexicográfico* con $x_1 > x_2 > x_3 > \dots > x_n$, donde

$$x_1^{\alpha_1}x_2^{\alpha_2} \dots x_n^{\alpha_n} \prec x_1^{\beta_1}x_2^{\beta_2} \dots x_n^{\beta_n} \text{ si } \alpha_1 < \beta_1 \text{ o } \alpha_1 = \beta_1 \text{ y } \alpha_2 < \beta_2 \text{ o } \dots$$

Por ejemplo, $x_1x_2 \prec x_1^3x_2$, $x_1x_3 \prec x_1x_2$.

Fijado un orden monomial, notamos con $Lt(f)$ al término principal de f (el término con mayor monomio).

Definición: Sean $\{f_1, \dots, f_s\} \subseteq k[x_1, \dots, x_n]$, k cuerpo. Se tiene un *algoritmo de división* con respecto a un orden fijo cuando se tiene un algoritmo tal que dado $f \in k[x_1, \dots, x_n]$ produce $q_1, \dots, q_s, r \in k[x_1, \dots, x_n]$ tales que

$$f = q_1f_1 + \dots + q_sf_s + r$$

con

1. Si $q_i \neq 0$, $Lt(q_if_i) \preceq Lt(f)$.
2. Ningún monomio de r es divisible por $Lt(f_1), \dots, Lt(f_s)$.

Llamamos $r_{\{f_1, \dots, f_s\}}(f)$ a r .

Ejemplo:

$$\begin{array}{r} x_1^2x_2 \quad -x_2^3 \quad +x_2x_3 \\ - \\ x_1^2x_2 \quad +x_1x_2^2 \\ \hline -x_1x_2^2 \quad -x_2^3 \quad +x_2x_3 \\ - \\ -x_1x_2^2 \quad -x_2^3 \\ \hline \quad \quad \quad x_2x_3 \\ - \\ \quad \quad \quad x_2x_3 \quad -x_2x_4 \\ \hline \quad \quad \quad \quad \quad x_2x_4 \\ \hline \end{array} \begin{array}{l} \frac{x_1 + x_2}{x_1x_2 - x_2^2} \\ \\ \\ \frac{x_3 - x_4}{x_2} \\ \\ \end{array}$$

$$\Rightarrow x_1^2x_2 - x_2^3 + x_2x_3 = (x_1x_2 - x_2^2)(x_1 + x_2) + x_2(x_3 - x_4) + x_2x_4$$

donde $Lt((x_1x_2 - x_2^2)(x_1 + x_2)) = x_1^2x_2 = Lt(x_1^2x_2 - x_2^3 + x_2x_3)$, $Lt(x_2(x_3 - x_4)) = x_2x_3 \preceq Lt(x_1^2x_2 - x_2^3 + x_2x_3)$ y $x_1 \nmid x_2x_4$, $x_3 \nmid x_2x_4$.

Observación: Si dividimos f por $\{f_1, \dots, f_s\}$ y obtenemos $f = q_1f_1 + \dots + q_sf_s + r$, entonces,

$$f \in \langle f_1, \dots, f_s \rangle \Leftrightarrow r_{\{f_1, \dots, f_s\}} \in \langle f_1, \dots, f_s \rangle.$$

Definición: Sea I ideal de $k[x_1, \dots, x_n]$. Fijado un orden monomial, se dice que $\{g_1, \dots, g_t\}$ es una *base de Gröbner* de I para dicho orden si

- $g_1, \dots, g_t \in I$ y,
- vale para cualquier $f \in k[x_1, \dots, x_n]$ que

$$f \in I \Leftrightarrow r_{\{g_1, \dots, g_t\}}(f) = 0.$$

Ejercicios:

1) Analicemos el siguiente código de Singular:

```

ring r = (0,k1,k2,k3,k4,k5,k6,C1,C2,C3), (x6,x5,x1,x3,x4,x2), lp;
poly f1 = -k1*x1*x3+k2*x5+k6*x6;
poly f2 = -k4*x2*x4+k5*x6+k3*x5;
poly f3 = -k1*x1*x3+(k2+k3)*x5;
poly f4 = -k4*x2*x4+(k5+k6)*x6;
poly f5 = -f3;
poly f6 = -f4;
poly g1 = x3+x5-C1;
poly g2 = x4+x6-C2;
poly g3 = x1+x2+x5+x6-C3;
ideal i1 = f1,f3,f4;
print("");
eliminate(i1,x5*x6);
print("");

```

declaramos los polinomios f_1, f_2, \dots, f_6

declaramos un anillo de característica 0, con parámetros $k_1, \dots, k_6, C_1, C_2, C_3$, variables $x_6 > x_5 > x_1 > x_3 > x_4 > x_2$ y orden lexicográfico

declaramos las leyes de conservación

declaramos el ideal generado por f_1, f_3, f_4

le pedimos que nos muestre el resultado de eliminar x_5 y x_6 operando con f_1, f_3 y f_4

La salida de este programa es:

```

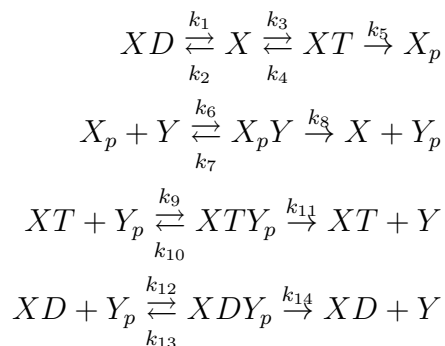
SINGULAR /
A Computer Algebra System for Polynomial Computations / version 4.1.1
0<
by: W. Decker, G.-M. Greuel, G. Pfister, H. Schoenemann \ Feb 2018
FB Mathematik der Universitaet, D-67653 Kaiserslautern \

_[1]=(k1*k3*k5+k1*k3*k6)*x1*x3+(-k2*k4*k6-k3*k4*k6)*x4*x2
. Auf Wiedersehen.

```

polinomio que se obtiene al eliminar x_5 y x_6 de f_1, f_3 y f_4

2) Consideremos la siguiente red que corresponde al sistema EnvZ-OmpR de *Escherichia coli* que consiste en la quinasa sensora EnvZ, y el regulador de respuesta OmpR. Este sistema de señalización es un sistema de dos componentes típico y es el Ejemplo (S60) del Material Suplementario Online del artículo de Shinar y Feinberg [6]:



Denotamos:

$$[XD] = x_1, [X] = x_2, [XT] = x_3, [X_p] = x_4, \\ [Y] = x_5, [X_pY] = x_6, [Y_p] = x_7, [XTY_p] = x_8, [XDY_p] = x_9 .$$

Calcular la base de Gröbner reducida de ideal generado por $\dot{x}_1, \dots, \dot{x}_9$, con respecto al orden lexicográfico $x_1 > x_2 > x_4 > x_5 > x_6 > x_8 > x_9 > x_3 > x_7$. Considerar los siguientes comandos:

```
option(redSB);
ideal i = f1,f2,f3,f4,f5,f6,f7,f8,f9;
ideal j = std(i);
j;
```

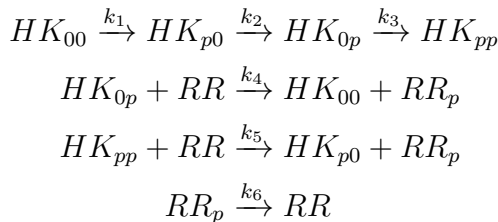
¿Cuánto vale x_7 en los estados estacionarios positivos? (Notar que no se consideraron las leyes de conservación para estos cálculos.)

- ¿Se puede obtener este valor operando linealmente sobre las ecuaciones originales?
- ¿Cómo se puede operar linealmente sobre las ecuaciones originales utilizando bases de Gröbner? (*Sugerencia: Pensar qué quiere decir "linealmente".*)

Fijadas ciertas cantidades totales $X_{tot} = [XD] + [X] + [XT] + [X_p] + [X_pY] + [XTY_p] + [XDY_p]$, $Y_{tot} = [Y] + [X_pY] + [Y_p] + [XTY_p] + [XDY_p]$, ¿cuántos estados estacionarios hay que satisfagan también estas dos últimas ecuaciones?

- ¿Cuántos estados estacionarios positivos hay por cada clase de compatibilidad estequiométrica?
- ¿Cuántos hay en el borde de cada clase de compatibilidad estequiométrica?

- 3) Consideremos ahora la misma red, pero agregando dos reacciones: la reacción reversa $X_pY \xrightarrow{k_{15}} X + Y_p$, y la reacción $Y_p \xrightarrow{k_{16}} Y$ [4]. Calcular la base de Gröbner reducida del ideal generado por $\dot{x}_1, \dots, \dot{x}_9$, con respecto al orden lexicográfico $x_1 > x_2 > x_4 > x_5 > x_6 > x_8 > x_9 > x_3 > x_7$. Deducir una cota superior para $[Y_p]$ en equilibrio. Si $[XT]$ en equilibrio es positiva, conseguir una cota que solo dependa de las constantes de reacción.
- 4) Consideremos la siguiente red que corresponde a otro sistema de dos componentes con histidina kinasa *híbrida*, estudiado en [5]. La red es la siguiente:

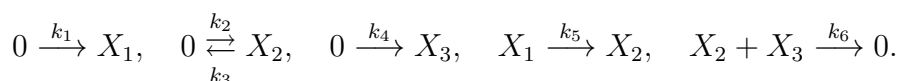


Llamamos $x_1 = [HK_{00}]$, $x_2 = [HK_{p0}]$, $x_3 = [HK_{0p}]$, $x_4 = [HK_{pp}]$, $x_5 = [RR]$ y $x_6 = [RR_p]$. Tenemos dos leyes de conservación

$$\begin{aligned}x_1 + x_2 + x_3 + x_4 &= C_1, \\x_5 + x_6 &= C_2.\end{aligned}$$

- Calcular una base de Gröbner reducida del ideal generado por $\dot{x}_1, \dot{x}_2, \dot{x}_3, \dot{x}_5, x_1 + x_2 + x_3 + x_4 - C_1$ y $x_5 + x_6 - C_2$ con respecto al orden lexicográfico $x_1 > x_2 > x_3 > x_4 > x_6 > x_5$. Observar que a lo sumo hay 3 estados estacionarios para cualquier valor de las constantes de reacción y totales de conservación.
- Observar que las ecuaciones correspondientes a $\dot{x}_1, \dot{x}_2, \dot{x}_3, \dot{x}_5, x_1 + x_2 + x_3 + x_4 - C_1$ son lineales en las variables x_1, \dots, x_4, x_6 . Considerar x_5 como parámetro (¿qué pasa si $x_5 = 0$?) y despejar x_1, \dots, x_4, x_6 en función de x_5, k_1, \dots, k_6 y C_1 . Notar que si x_5 es positivo, entonces x_1, \dots, x_4, x_6 también lo son.
- Trabajar con el polinomio $p(x_5)$ en x_5 obtenido en el ítem 1, para encontrar parámetros k, C , tales que la red admita 3 estados estacionarios positivos. (*Sugerencia 1:* dado un polinomio $q(x_5)$ de grado 3 con tres raíces positivas, por ejemplo $q(x_5) = (x_5 - 1)(x_5 - 2)(x_5 - 3)$, podemos ajustar los valores de k, C de forma tal que $p(x_5) = q(x_5)$). *Sugerencia 2:* Se pueden tomar, por ejemplo, $k_3 = k_5 = k_6 = 1$, $C_2 = 9$ y despejar los otros parámetros usando el comando `solve` de la librería `solve.lib`).

5) Consideremos el siguiente sistema:



Las ecuaciones de este sistema son:

$$\begin{aligned}\dot{x}_1 &= k_1 - k_5 x_1, \\ \dot{x}_2 &= k_2 - k_3 x_2 + k_5 x_1 - k_6 x_2 x_3, \\ \dot{x}_3 &= k_4 - k_6 x_2 x_3.\end{aligned}$$

Si llamamos $g(t) = x_1(t) + x_2(t) - x_3(t)$, entonces $\frac{dg}{dt}(t) = k_1 + k_2 - k_4 - k_3 x_2(t)$. Notemos que si $k_1 + k_2 - k_4 < 0$, entonces $\frac{dg}{dt}(t) \leq k_1 + k_2 - k_4 < 0$ para todo $t \geq 0$ y g es estrictamente decreciente. Más aún, si x_1, x_2 y x_3 están acotadas para todo tiempo $t \geq 0$, entonces g converge decrecientemente a un límite y su derivada tiende a cero, lo que lleva a una contradicción pues $\frac{dg}{dt}(t) \leq k_1 + k_2 - k_4 < 0$. ¿Qué podemos concluir entonces? En otro sistema, ¿cómo podríamos obtener una función como esta g , que nos permita sacar una conclusión equivalente?

Referencias

- [1] COX D., LITTLE J., O'SHEA D., *Ideals, varieties, and algorithms: an introduction to computational algebraic geometry and commutative algebra*, Springer-Verlag, New York (1992).
- [2] COX D., LITTLE J., O'SHEA D., *Using algebraic geometry*, Graduate Texts in Mathematics, Vol.185, Second Edition, Springer, New York, (2005).
- [3] DECKER, W.; GREUEL, G.-M.; PFISTER, G.; SCHÖNEMANN, H., SINGULAR 4-1-1 — A computer algebra system for polynomial computations. <http://www.singular.uni-kl.de> (2018).
- [4] KARP R., PÉREZ MILLÁN M., DASGUPTA T., DICKENSTEIN A., GUNAWARDENA J., *Complex-linear invariants of biochemical networks*, J. Theor. Biol., Vol. 311, 130–138, (2012).
- [5] KOTHAMANCHU V. B., FELIU E., CARDELLI L., SOYER O. S., *Unlimited multistability and Boolean logic in microbial signaling*, J. R. Soc. Interface, 12 20150234 (2015).
- [6] G. SHINAR, AND M. FEINBERG, *Structural sources of robustness in biochemical reaction networks*, Science 327(5971), 1389–1391, (2010).