SOLVING SYSTEMS OF MULTIVARIATE POLYNOMIAL EQUATIONS BY EIGENVECTORS

- Solving polynomial systems using lexicographic GB and back-substitution may be time-consuming and numerically unstable (leads to large matrices M_i)

- An alternative approach allows to compute all solutions at once as eigenvector problem

Example using one polynomial in one unknown

$$f = x^3 - 6x^2 + 11x - 6 = (x-1)(x-2)(x-3) = 0$$

The roots can be find as eigenvalues of the companion matrix

$$M_x = \begin{pmatrix} 0 & 0 & 6 \\ 1 & 0 & -11 \\ 0 & 1 & 6 \end{pmatrix}$$

- Roots can be found by computing eigenvectors of M_x^T

- Let's consider remainders of all polynomials $g \in \mathbb{Q}[x]$ on divison by f
  - It is the set of all polynomials r of degree at most 2
  - All polynomials of degree at most 2 are left unchanged by the long division by f and all monomials of a higher degree will get rewritten using f in terms of polynomials of degree at most 2

- We can write:

$$r = a_2x^2 + a_1x + a_0 \quad , \quad a_0, a_1, a_2 \in \mathbb{Q}$$

  ⇒ We can identify each reminder r with a 3-dimensional vector $r \equiv [a_0, a_1, a_2]^T \in \mathbb{Q}^3$

  ⇒ the set of all such reminders is in one-to-one correspondence with $\mathbb{Q}^3$

- Now consider the mapping $M_x : \mathbb{Q}[x] \to \mathbb{Q}[x]$ on polynomials given by

$$M_x(h) = (xh) \bmod f$$

- It maps polynomials of degree at most 2 back to polynomials of degree at most 2

$$M_x(1) = x.1 \bmod f = x \bmod f = x$$
$$M_x(x) = x.x \bmod f = x^2 \bmod f = x^2$$
$$M_x(x^2) = x.x^2 \bmod f = x^3 \bmod f = 6x^2 - 11x + 6$$

- M_x is a linear mapping since for all $g, h \in \mathbb{Q}[x]$, $a \in \mathbb{Q}$ we have

$$M_x(g+h) = (x.g + x.h) \bmod f = (xg) \bmod f + (xh) \bmod f = M_x(g) + M_x(h)$$
$$M_x(a.g) = (a.x.g) \bmod f = a(xg) \bmod f = a M_x(g)$$

$\Rightarrow$ M_x is a linear mapping on the set of all polynomials of degree 2

$$M_x(a_2x^2 + a_1x + a_0) = a_2 M_x(x^2) + a_1 M_x(x) + a_0 M_x(1)$$

- Every linear mapping has a matrix of the mapping w.r.t. a fixed basis

- Let us choose the standard monomial basis $[1, x, x^2]$ in the linear space of $\mathbb{Q}[x]$ of polynomials of degree at most 2 and write the above represented by vectors in $\mathbb{Q}^3$.
We will express monomials as vectors using the basis $[1\ x\ x^2]$

$$M_x(1) \equiv M_x\left(\begin{bmatrix}1\\0\\0\end{bmatrix}\right) = \begin{bmatrix}0\\1\\0\end{bmatrix}$$

$$M_x(x) \equiv M_x\left(\begin{bmatrix}0\\1\\0\end{bmatrix}\right) = \begin{bmatrix}0\\0\\1\end{bmatrix}$$

$$M_x(x^2) \equiv M_x\left(\begin{bmatrix}0\\0\\1\end{bmatrix}\right) = \begin{bmatrix}6\\-11\\6\end{bmatrix}$$

To get the matrix of the mapping M_x we write

$$M_x\left(\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}\right) = \begin{bmatrix}0&0&6\\1&0&-11\\0&1&6\end{bmatrix} = \mathbf{M}_x$$

$\Rightarrow$ matrix $\mathbf{M}_x$ of M_x w.r.t. the standard monomial basis is the companion matrix

Let us evaluate polynomials $g \in \mathbb{Q}[x]$ on the roots of f.

- Consider a root p of f, i.e a solution to the equation $f(x) = 0$ ($f(p) = 0$)

- In our example $f(x) = x^3 - 6x^2 + 11x - 6$ we have 3 roots p_1, p_2, p_3
  Let us evaluate polynomials x, x^2, x^3 on the roots p_i:

$$X(p_i) = p_i = p_i \cdot 1 = p_i \cdot 1(p_i) = X(p_i) \cdot 1(p_i)$$
$$X^2(p_i) = p_i^2 = p_i \cdot p_i = p_i \cdot X(p_i) = X(p_i) \cdot X(p_i)$$
$$X^3(p_i) = p_i^3 = p_i \cdot p_i^2 = p_i \cdot X^2(p_i) = X(p_i) \cdot X^2(p_i)$$

Now since $X^3(p_i) = M_X(x^2)(p_i) = (6x^2 - 11x + 6)(p_i)$ we get

$$(6x^2 - 11x + 6)(p_i) = X(p_i)\, X^2(p_i)$$

$$(6x^2 - 11x + 6)(p_i) = x(p_i)\,x^2(p_i)$$

We can write

$$x(p_i)\begin{bmatrix} 1(p_i) \\ x(p_i) \\ x^2(p_i) \end{bmatrix} = \begin{bmatrix} x(p_i) \\ x^2(p_i) \\ x^3(p_i) \end{bmatrix} = \begin{bmatrix} x(p_i) \\ x^2(p_i) \\ (6x^2 - 11x + 6)(p_i) \end{bmatrix}$$

$$x(p_i)\begin{bmatrix} 1(p_i) \\ x(p_i) \\ x^2(p_i) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 6 & -11 & 6 \end{bmatrix}\begin{bmatrix} 1(p_i) \\ x(p_i) \\ x^2(p_i) \end{bmatrix}$$

$$x(p_i)\begin{bmatrix} 1 \\ p_i \\ p_i^2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 6 & -11 & 6 \end{bmatrix}\begin{bmatrix} 1 \\ p_i \\ p_i^2 \end{bmatrix}$$

$$\underline{p_i \cdot \vec{v}_i = M_x^T \vec{v}_i}$$

$\Rightarrow (p_i, \vec{v}_i)$ are eigenvalue-eigenvector pairs of M_x^T

- Eigenvalues p_i are evaluations of x on the roots of f and eigenvectors $\vec{v}_i$ are evaluations of the standard basis $[1\ x\ x^2]$ on the roots of f

- This observation holds true in general
- For a polynomial f of degree n we are getting a $n \times n$ matrix with n eigenvalues counting with multiplicities

- When the matrix M_x has separated one-dimensional eigenspaces, which happens always when eigenvalues are pairwise distinct, i.e when f has all roots with multiplicity one, we can compute basis w_i of each eigenspace and get v_i as

$$\vec{v_i} = \frac{1}{w_{in}} \vec{w_i} \qquad i = 1, \dots, n$$

- Solutions to f are obtained from $\vec{v_i}$ as $p_i = x(p_i) = v_{i2}$

- It is possible to generalize this to more general mapping

$$M_h : \mathbb{Q}[x] \to \mathbb{Q}[x], \quad M_h(g) = (h.g) \bmod f$$

by replacing x by a general polynomial $h \in \mathbb{Q}[x]$

The key concept for deriving the relationship between the solutions to $f(x)=0$ and the eigenvectors of M_h (M_x) in the univariate case was the remainder r of h on the division by f gave the values of h on the roots of f

$$h = q \cdot f + r \qquad h(p) = q(p)\underbrace{f(p)}_{=0} + r(p)$$
$$h(p) = r(p)$$

$\Rightarrow$ Long division produced $r = h - qf$ such that r was "the simplest" polynomial evaluation on the roots of f to the same values as h

- We could also see this as removing from h all what can be generated by f, i.e $\langle f \rangle = \{ g \cdot f \mid g \in \mathbb{Q}[x] \}$

- We can also say that r is equivalent to h writing
$$h \equiv r \quad \text{when} \quad h - r = q \cdot f \in I = \langle f \rangle$$

Solving systems of multivariate polynomial equation by eigenvectors

- Generalization to systems of p.e. in several unknowns

- In the multivariate case

$$I = \langle f_1, \dots, f_n \rangle = \{ \sum_{i=1}^{n} g_i f_i \mid g_i \in \mathbb{Q}[x_1, \dots, x_n] \}$$

- In the univariate case the remainders r on the long division by f had a good property that all monomials of r were strictly smaller (when ordered by the degree) than the LM of f

- The maximal degree of r was equal to the number of solutions - 1 (m-1) and r was a linear combination of exactly m monomials (counting $x^0 = 1$)

- That gave $m \times m$ multiplication matrix M_h

- This was thanks to the fact that the ideal $\langle f \rangle$ was in one-to-one correspondence with its generator f

- In the multivariate case $I = \langle f_1, \ldots, f_n \rangle$ can be generated by infinitely many different sets of generators and in general there is no direct connection between the multidegrees of the LMs of a particular generator set and the number of solutions

- Futhermove with a general set of generators F of I remainders on division by F are not well defined
  - different r's can be obtained when changing the order of f_j

- Fortunately for reduced GB^G we have a unique remainder r on division by G independently on the order in which are the generators G used in the division process

- Reminder $r = g \bmod_{<_o} G$ is thus defined uniquely by the ideal I and the monomial ordering $<_o$ used

- Further r is a linear combination of monomials that are not divisible by any leading monomial of generators G

- The actual monomials may by different dependent on the monomial ordering $<_o$ used, but their number l will be the same

- The relationship between l and the number of solutions m is in general

$$l \geq m$$

The equality occures exactly when I is a radical ideal

- Radical ideal - I is such that $f^h \in I$ for some h implies $f \in I$
- Intuitively radicality is connected to multiplicity of solutions
- Radical ideals have no multiplicities in any coordinate

Generalize the eigenvector method to polynomial system $F = \{f_1, \dots, f_n\}$ in n unknowns $x_1, \dots, x_n$

1. Fix a particular monomial ordering $<_0$
2. Construct the reduced GB G of $I = \langle F \rangle$ for $<_0$
3. Construct the set B of all monomials that are divisible by no leading monomial of all polynomials in G
4. Fix a polynomial $g \in \mathbb{Q}[x_1, \dots, x_n]$ such that g has different solutions e.g. take a random linear polynomial. This guarantees isolated one-dimensional eigenspaces for the radical ideal $\langle F \rangle$
5. Construct the multiplication matrix by finding remainders of $g \cdot b$ for all $b \in B$ on division by G w.r.t. $<_0$
6. Find eigenvalues and eigenvectors of M_g (for radical ideal eigenspaces are one-dimensional)
7. Recover the solutions from eigenvalues, eigenvectors and G

$<_0$ - good ordering Grevlex - is archimedean - there is only finitely many monomials smaller than any monomial

Consider a polynomial system $F = \{f_1, f_2\}$

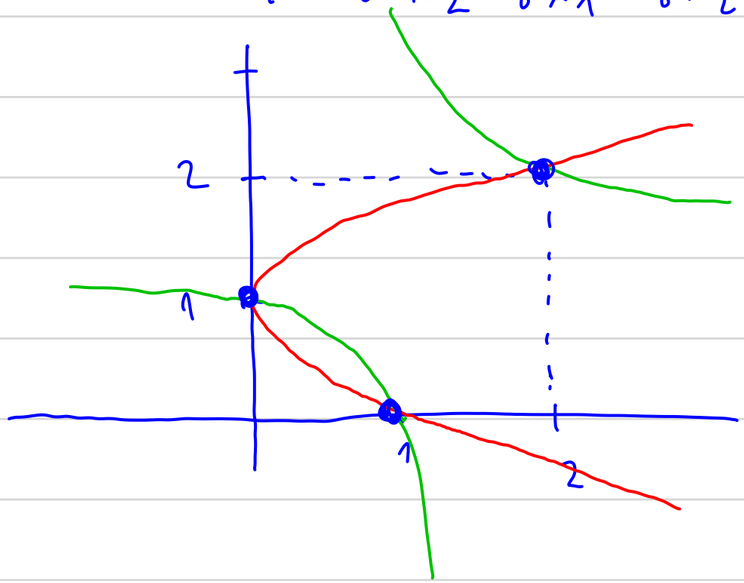
$$f_1 = 6x_1x_2 + 3x_2^2 - 10x_1 - 13x_2 + 10$$
$$f_2 = 3x_2^2 - 2x_1 - 5x_2 + 2$$

The system has 3 solutions, all with multiplicity one. Ideal $\langle F \rangle$ is radical

$$\prec_0 \equiv x_2 \prec_{\text{grevlex}} x_1$$

Monomials of F will be thus ordered as

$$1 \prec_0 x_2 \prec_0 x_1 \prec_0 x_2^2 \prec_0 x_1x_2$$



Solutions to 2 conics are

$[0,1]$, $[1,0]$, $[2,2]$

- To get an eigenvalue / eigenvector problem, we need to find a multiplication matrix for a polynomial w.r.t. that we will generate all remainders on the division by GB of $\langle F \rangle$

-

- With grevlex ordering we expect B to contain 3 smallest monomials $1, x_1, x_2$ (all remainders will be linear combinations of $1, x_1, x_2$)

$$f_1 = 6x_1x_2 + 3x_2^2 - 10x_1 - 13x_2 + 10$$

- Let's construct GB $\quad f_2 = 3x_2^2 - 2x_1 - 5x_2 + 2$

$$\mathrm{LCM}(x_1x_2, x_2^2) = x_1x_2^2$$

$$S(f_1, f_2) = \frac{x_1x_2^2}{6x_1x_2} f_1 - \frac{x_1x_2^2}{3x_2^2} f_2 = \frac{x_2}{6} f_1 - \frac{x_1}{3} f_2 = (3x_2^3 + 4x_1^2 - 13x_2^2 - 4x_1 + 10x_2)/6$$

$$f_3 = \overline{S(f_1, f_2)}^{F} = 3x_1^2 - 5x_1 - 2x_2 + 2 \qquad G = \{f_1, f_2\} \cup \{f_3\}$$

$$\overline{S(f_1, f_2)}^{G} = 0, \quad \overline{S(f_1, f_3)}^{G} = 0, \quad \overline{S(f_2, f_3)}^{G} = 0$$

$\Rightarrow$ No new non-zero remainder has been generated $\Rightarrow$ we have obtained GB of $\langle F \rangle$

- We can simplify G to obtain reduced GB of $\langle F \rangle$
- The idea is to remove all monomials from polynomials of G that can be divided by the LT of G
  - it is a generalization of G-J elimination

- In this case there is monomial x_2^2 in f_1 that is divisible by the leading term x_2^2 of f_2, hence we can remove it by subtracting f_2 from f_1

The reduced GB is

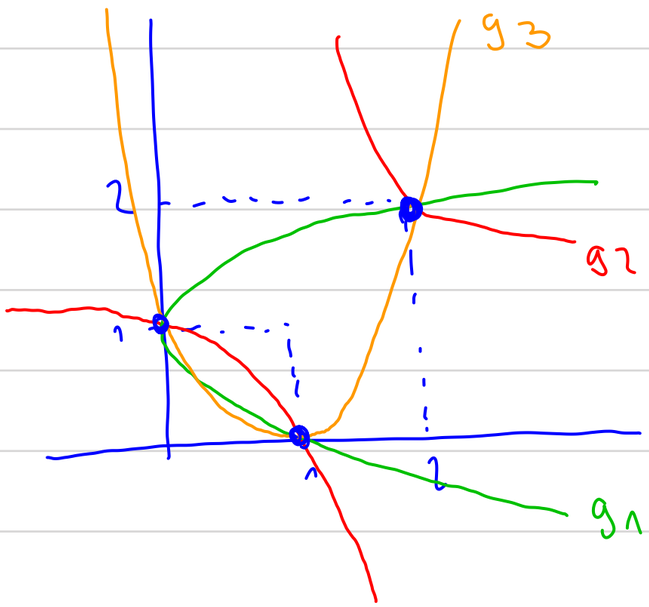
$$g_1 = x_1 x_2 - \frac{4}{3} x_1 - \frac{4}{3} x_2 + \frac{4}{3}$$

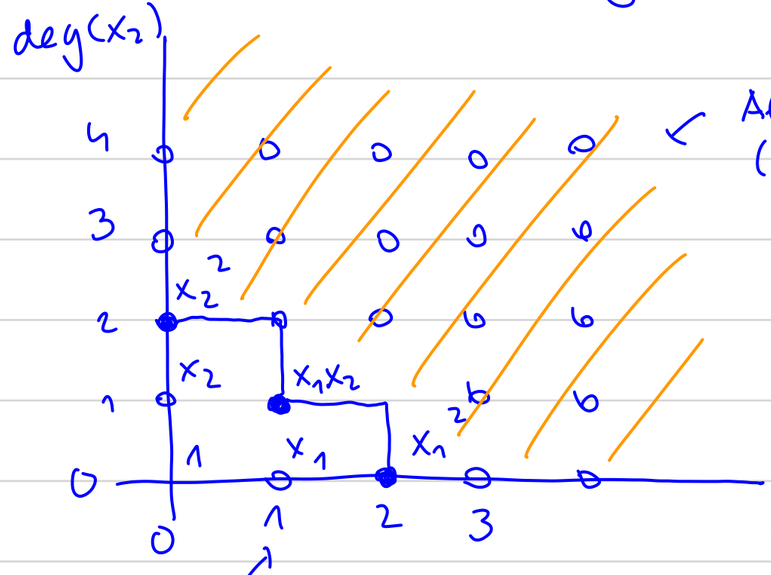
$$g_2 = x_2^2 - \frac{2}{3} x_1 - \frac{5}{3} x_2 + \frac{2}{3}$$

$$g_3 = x_1^2 - \frac{5}{3} x_1 - \frac{2}{3} x_2 + \frac{2}{3}$$

- The leading monomials of G_r, i.e. x_1x_2, x_2^2 and x_1^2 reduce all monomials except for the three monomials $1, x_1, x_2$

- These are the three monomials that will provide the basis of the linear space to form a multiplication matrix and to obtain eigenvalue / eigenvector problem providing us with the solution to the original system F





standard monomials $x_1, x_2, 1$ of G are not divisible by a leading monomials x_1^2, x_1x_2, x_2^2 of G

Let's consider mapping $M_g : \mathbb{Q}[x_1, x_2] \to \mathbb{Q}[x_1, x_2]$ by a polynomial $g \in \mathbb{Q}[x_1, x_2]$ defined by

$$M_g(h) = \overline{(g.h)}^G \qquad G\text{-GB of } F$$

The reduction of $g.h$ as well as the computation of G is carried out w.r.t the same monomial ordering

- Matrices M_{x_1}, M_{x_2} $(g = x_1 / g = x_2)$ can be extracted from G_r

We have

$$\begin{array}{c} \\ g_1 \\ g_2 \\ g_3 \end{array} \begin{array}{c} \begin{array}{cccccc} x_1^2 & x_1x_2 & x_2^2 & x_1 & x_2 & 1 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & -\frac{5}{3} & -\frac{2}{3} & \frac{2}{3} \\ 0 & 1 & 0 & -\frac{4}{3} & -\frac{4}{3} & \frac{4}{3} \\ 0 & 0 & 1 & -\frac{2}{3} & -\frac{5}{3} & \frac{2}{3} \end{bmatrix} \end{array}$$

We see

$$x_1 \begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix} \overset{G}{=} \begin{bmatrix} x_1^2 \\ x_1x_2 \\ x_1 \end{bmatrix} \overset{G}{=} \begin{bmatrix} \frac{5}{3} & \frac{2}{3} & -\frac{2}{3} \\ \frac{4}{3} & \frac{4}{3} & -\frac{4}{3} \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix}$$

$M_{x_1}^T$

$$x_2 \begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix} \overset{G}{=} \begin{bmatrix} x_1x_2 \\ x_2^2 \\ x_2 \end{bmatrix} \overset{G}{=} \begin{bmatrix} \frac{4}{3} & \frac{4}{3} & -\frac{4}{3} \\ \frac{2}{3} & \frac{5}{3} & -\frac{2}{3} \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix}$$

$M_{x_2}^T$

Since the system F has 3 solutions with multiplicity one, $\langle F \rangle$ is radical

$$[1,0], [0,1], [2,2]$$

Since all three solutions have pairwise distinct x_1 (as well as x_2) $(0,1,2)$ we can choose $g = x_1$ and then $M_g = M_{x_1}$

- We compute eigenvectors of $M_{x_1}^T$ and get 3 one-dimensional bases of 3 one-dimensional eigenspaces

$$\text{eigvect}(M_{x_1}^T) = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix}$$ ← corresponding to evaluations of $\begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix}$ on solutions p_1, p_2, p_3

⇒ we get 3 solutions $[1,0], [0,1], [2,2]$