Solving systems of multivariate polynomial equations by eigenvectors

- Solving polynomial systems using lexicographic GB and Buchberger may be time-consuming and numerically unstable (leads to large matrices $\Pi_i$.)

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

Example using one polynomial in one unknown

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

The roots can be found as eigenvalues of the companion matrix

$\Pi x = \begin{pmatrix} 0 & 0 & 6 \\ 1 & 0 & -11 \\ 0 & 1 & 6 \end{pmatrix}$
- Roots can be found by computing eigenvectors of $M_x^T$

- Let's consider remainders of all polynomials $g \in \mathbb{Q}[x]$ on division 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:

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

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

- The set of all such remainders 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
  $M_x(h) = (xh) \mod f$

- It maps polynomials of degree at most 2 back to polynomials of degree at most 2:
  $M_x(1) = x \cdot 1 \mod f = x \mod f = x$
  $M_x(x) = x \cdot x \mod f = x^2 \mod f = x^2$
  $M_x(x^2) = x \cdot x^2 \mod f = x^3 \mod f = 6x^2 - M_x(x) + 6$

- $M_x$ is a linear mapping since for all $g, h \in \mathbb{Q}[x], a \in \mathbb{Q}$ we have:
  $M_x(g + h) = (x \cdot g + x \cdot h) \mod f = (xg) \mod f + (xh) \mod f = M_x(g) + M_x(h)$
  $M_x(ah) = (a \cdot x \cdot g) \mod f = a(xg) \mod f = a \cdot M_x(g)$

$\Rightarrow$ $M_x$ 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 \([1, x, x^2]\)

\[
M_x(1) = M_x \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}
\]

\[
M_x(x) = M_x \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}
\]

\[
M_x(x^2) = M_x \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}
\]

To get the matrix of the mapping \(M_x\) we write

\[
M_x \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = M_x
\]

\(\Rightarrow\) matrix \(M_x\) 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 \( f(x) = x^3 - 6x^2 + Mx - 6 \) we have 3 roots \( p_1, p_2, p_3 \).

Let us evaluate polynomials \( x, x^2, x^3 \) on the roots \( p_i \):

\[
\begin{align*}
x(p_i) &= p_i = p_i, \\
x^2(p_i) &= p_i^2 = p_i, \\
x^3(p_i) &= p_i^3 = p_i.
\end{align*}
\]

\[
\begin{align*}
\lambda(p_i) &= p_i, \\
\lambda(p_i)^2 &= p_i, \\
\lambda(p_i)^3 &= p_i.
\end{align*}
\]

Now since \( x^3(p_i) = M_x(x^2)(p_i) = (6x^2 - Mx + 6)(p_i) \) we get

\[
(6x^2 - Mx + 6)(p_i) = x(p_i) x^2(p_i).
\]
We can write

\[(6x^2 - 11x + 6)(p_i) = x(p_i)x^2(p_i)\]

\[
\begin{bmatrix}
1(p_i) \\
x(p_i) \\
x^2(p_i)
\end{bmatrix}
= 
\begin{bmatrix}
x(p_i) \\
x^2(p_i) \\
x^3(p_i)
\end{bmatrix}
= 
\begin{bmatrix}
x(p_i) \\
x^2(p_i)
\end{bmatrix}
\]

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

\[
\begin{bmatrix}
\Lambda(p_i) \\
\pi(p_i) \\
\pi^2(p_i)
\end{bmatrix}
= 
\begin{bmatrix}
0 & 1 & 0 \\
0 & 0 & 1 \\
6 & -11 & 6
\end{bmatrix}
\begin{bmatrix}
\Lambda(p_i) \\
\pi(p_i) \\
\pi^2(p_i)
\end{bmatrix}
\]

\[
p_i \cdot \vec{\nu}_i = M_x^T \vec{\nu}_i
\]

\[
\Rightarrow (p_i, \vec{\nu}_i) \text{ are eigenvalue-eigenvector pairs of } M_x^T
\]

- Eigenvalues \(p_i\) are evaluations of \(x\) on the roots of \(f\) and eigenvectors \(\vec{\nu}_i\) 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 $v_i$ of each eigenspace and get $v_i$ as

$$\vec{v}_i = \frac{1}{\mu_i} \vec{w}_i, \quad i = 1, ..., n$$

- 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

$$M_{a, b} : \mathbb{Q}[x] \to \mathbb{Q}[x], \quad M_{a, b}(g) = (a \cdot g) \mod f$$

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 \( M_\alpha \) (\( M_x \)) in the univariate case was the remainder \( r \) of \( h \) on the division by \( f \) gave the values of \( h \) on the roots of \( f \).

\[
\begin{align*}
\hat{h} &= q \cdot f + r \\
\hat{h}(p) &= q(p)f(p) + r(p) \\
h(p) &= r(p)
\end{align*}
\]

Long division produced \( r = h - qf \) such that \( r \) was “the simplest” polynomial evaluating 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. \( \langle f \rangle = \{ g \cdot f \mid g \in \mathbb{C}[x] \} \).

We can also say that \( r \) is equivalent to \( h \) writing \( h \equiv r \) when \( h - r = q \cdot f \in I = \langle f \rangle \).
Solving systems of multivariate polynomial equation by eigenvectors
- Generalization to systems of p.e. in several unknowns

- In the multivariate case

\[ I = \langle f_n, \ldots, f_1 \rangle = \{ \sum g_i f_i : g_i \in \mathbb{Q}[x_1, \ldots, x_n] \} \]

- 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 \( x^n \))

- That gave \( m \times m \) multiplication matrix \( M_r \)

- 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 $I = \langle f_1, \ldots, f_k \rangle$ 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.

- Furthermore, 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 $f_j$.

\[ g \]

- Fortunately for reduced GB's 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.

- Remainder $r = g \mod \subseteq_0 G$ is thus defined uniquely by the ideal $I$ and the monomial ordering $\subseteq_0$ used.
- Further \( r \) is a linear combination of monomials that are not divisible by any leading monomial of generators \( G \).

- The actual monomials may be different dependent on the monomial ordering \( \preceq \) 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 occurs exactly when \( I \) is a radical ideal.

  - Radical ideal - \( I \) is such that \( f^k \in I \) for some \( k \) 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 \( F = \{ f_1, \ldots, f_n \} \) in \( n \) unknowns \( x_1, \ldots, x_n \)

1. Fix a particular monomial ordering \( \leq \).
2. Construct the reduced GB \( G \) of \( I = \langle F \rangle \) for \( \leq \).
3. Construct the set \( B \) of all monomials that are divisible by no leading monomial of all polynomials in \( G \).
4. Fix a polynomial \( g \in \mathbb{Q}[x_1, \ldots, x_n] \) 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. \( \leq \).
6. Find eigenvalues and eigenvectors of \( Mg \) (for radical ideal eigenspaces are one-dimensional).
7. Recover the solutions from eigenvalues, eigenvectors and \( G \).

\( \leq \) - 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

\[
\langle 0 \rangle \equiv X_2 \leq \text{gre藤} \equiv X_1
\]

Monomials of $F$ will be thus ordered as

\[
1 \leq_0 x_2 \leq_0 x_1 \leq_0 x_2^2 \leq_0 x_1x_2
\]

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 \( 1, x_1, x_2 \)).

\[
\begin{align*}
  f_1 &= 6x_1x_2 + 3x_2^2 - 10x_1 - 13x_2 + 16 \\
  f_2 &= 3x_1^2 - 2x_1 - 5x_2 + 2 \\
  \text{LCM} (x_1x_2, x_2) &= x_1x_2^2 \\
  S(t_1, t_2) &= \frac{x_1x_2}{6x_1}, f_1 - \frac{x_1x_2}{2x_1} f_2 = \frac{x_2}{6} f_1 - \frac{x_4}{3} f_2 = \left(3x_2^3 + 4x_1x_2 - 13x_2^2 - 4x_1 + 10x_2 \right) / 6 \\
  f_3 &= S(t_1, t_2) = 3x_1^2 - 5x_1 - 2x_2 + 2 \quad G = \{ t_1, t_2 \} \cup \{ t_3 \} \\
  S(t_1, t_2) &= 0 \quad S(t_1, t_3) = 0 \quad S(t_2, t_3) = 0
\end{align*}
\]

\( \Rightarrow \) 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 $x_2^3$ in $f_1$ that is divisible by the leading term $x_2$ of $f_2$, hence we can remove it by subtracting $f_2$ from $f_1$.

The reduced GB is

\[
\begin{align*}
g_1 &= x_1x_2 - \frac{4}{3}x_1 - \frac{5}{6}x_2 + \frac{5}{3} \\
g_2 &= x_2^2 - \frac{2}{3}x_1 - \frac{5}{3}x_2 + \frac{2}{3} \\
g_3 &= x_1^2 - \frac{5}{3}x_1 - \frac{2}{3}x_2 + \frac{2}{3}
\end{align*}
\]
- The leading monomials of $Gr$, i.e. $x_1 x_2$, $x_2^2$ 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 $F$.

\begin{align*}
\text{Standard monomials } x_1, x_2, 1 \text{ of } 6 \text{ are not divisible by a leading monomials } x_1^2, x_1 x_2, x_2^2 \text{ of } 6
\end{align*}
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

$$M_g (h) = \overline{g \cdot h} \quad G \cdot GB \text{ of } F$$

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

- Matrices $M_{x_1}, M_{x_2}$ ($y = x_1, y = x_2$) can be extracted from $G$

We have

$$g_1 = \begin{bmatrix} 1 & 0 & 0 & -\frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\
1 & 1 & 0 & -\frac{1}{3} & -\frac{1}{3} & \frac{5}{3} \\
0 & 0 & 1 & -\frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

We see

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

$$x_2 \begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix} = \overline{x_1 x_2} \begin{bmatrix} x_1^3 \\ x_1^2 x_2 \\ x_1 x_2^2 \\ x_2^3 \\ x_2^2 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{5}{3} & \frac{5}{3} & -\frac{1}{3} \\ \frac{2}{3} & \frac{4}{3} & -\frac{1}{3} \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix}$$
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 \( M_g = Mx_1 \)

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

\[ \text{eigvec} (M_{x_1}^T) = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix} \]

\( \text{corresponding to evaluations of} \)

\[ \begin{bmatrix} x_1 \\ x_2 \\ \lambda \end{bmatrix} \]

on solutions \( p_1, p_2, p_3 \)

\( \Rightarrow \) we get 3 solutions \( [1,0], [0,1], [2,2] \)