9. POLYNOMIAL EQUATIONS

- Laguerre's method
- Sturm sequences
- Resultants

Class web site:

http://www.di.ens.fr/~brette/calculscientifique/index.htm

http://www.di.ens.fr/~ponce/scicomp/notes.pdf

Newton's Method: Multivariate Case

What are the solutions of $f_1(\vec{x}) = \ldots = f_p(\vec{x}) = 0$?

- Taylor expansion of f_i in the neighborhood of x:

$$\begin{aligned} f_i(\vec{x}+\delta\vec{x}) &= f_i(\vec{x}) + \delta x_1 \frac{\partial f_i}{\partial x_1}(\vec{x}) + \ldots + \delta x_q \frac{\partial f_i}{\partial x_q}(\vec{x}) + O(|\delta\vec{x}|^2) \\ &\approx f_i(\vec{x}) + \nabla f_i(\vec{x}) \cdot \delta\vec{x}, \end{aligned}$$

where $\nabla f_i(\vec{x}) = (\partial f_i/\partial x_1, \ldots, \partial f_i/\partial x_q)^T$ is the *gradient* of f_i at the point $\vec{x}$

- This can be rewritten as

$$\vec{f}(\vec{x}+\delta\vec{x}) \approx \vec{f}(\vec{x}) + J_{\vec{f}}(\vec{x})\delta\vec{x} = \vec{0},$$

where $J_{\vec{f}}(\vec{x})$ is the *Jacobian* of $\vec{f} = (f_1, \ldots, f_n)^T$—that is, the $p \times q$ matrix

$$J_{\vec{f}}(\vec{x}) \stackrel{\text{def}}{=} \begin{bmatrix} \nabla f_1^T(\vec{x}) \\ \ldots \\ \nabla f_p^T(\vec{x}) \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\vec{x}) & \ldots & \frac{\partial f_1}{\partial x_q}(\vec{x}) \\ \ldots & \ldots & \ldots \\ \frac{\partial f_p}{\partial x_1}(\vec{x}) & \ldots & \frac{\partial f_p}{\partial x_q}(\vec{x}) \end{bmatrix}.$$

- **When** $p = q$: Iterate

$$\vec{x} = \vec{x} - J_{\vec{f}}^{-1}(\vec{x})\, \vec{f}(\vec{x}).$$

- Quadratic convergence rate when it converges.

Newton's Method for Nonlinear Least Squares

When $p > q$, define

$$E(\vec{x}) \stackrel{\text{def}}{=} \frac{1}{2}|\vec{f}(\vec{x})|^2 = \frac{1}{2}\sum_{i=1}^{p} f_i^2(\vec{x}),$$

and use Newton's method to find a local minimum of E as a zero of its gradient $\vec{F}(\vec{x}) = \nabla E(\vec{x})$.

• A simple calculation shows that

$$\vec{F}(\vec{x}) = J_{\vec{f}}^T(\vec{x})\vec{f}(\vec{x}),$$

and differentiating this expression shows in turn that the Jacobian of $\vec{F}$ is

$$J_{\vec{F}}(\vec{x}) = J_{\vec{f}}^T(\vec{x})J_{\vec{f}}(\vec{x}) + \sum_{i=1}^{p} f_i(\vec{x})H_{f_i}(\vec{x}),$$

where $H_{f_i}(\vec{x})$ denotes the *Hessian* of f_i—that is, the $q \times q$ matrix of second derivatives

$$H_{f_i}(\vec{x}) \stackrel{\text{def}}{=} \begin{bmatrix} \frac{\partial^2 f_i}{\partial x_1^2}(\vec{x}) & \dots & \frac{\partial^2 f_i}{\partial x_1 x_q}(\vec{x}) \\ \dots & \dots & \dots \\ \frac{\partial^2 f_i}{\partial x_1 x_q}(\vec{x}) & \dots & \frac{\partial^2 f_i}{\partial x_q^2}(\vec{x}) \end{bmatrix}.$$

• The term $\delta\vec{x}$ in Newton's method satisfies $J_{\vec{F}}(\vec{x})\delta\vec{x} = -\vec{F}(\vec{x})$. Equivalently, $\delta\vec{x}$ is the solution of

$$[J_{\vec{f}}^T(\vec{x})J_{\vec{f}}(\vec{x}) + \sum_{i=1}^{p} f_i(\vec{x})H_{f_i}(\vec{x})]\delta\vec{x} = -J_{\vec{f}}^T(\vec{x})\vec{f}(\vec{x}).$$

Variants of Newton's Method for Nonlinear Least Squares

Gauss-Newton. What is the value of $\delta\vec{x}$ that minimizes $E(\vec{x}+\delta\vec{x})$ for a given value of $\vec{x}$?

$$E(\vec{x}+\delta\vec{x}) = |\vec{f}(\vec{x}+\delta\vec{x})|^2 \approx |\vec{f}(\vec{x}) + J_{\vec{f}}(\vec{x})\delta\vec{x}|^2.$$

• The adjustment $\delta\vec{x}$ can be computed as the solution of $J_{\vec{f}}^{\dagger}(\vec{x})\delta\vec{x} = -\vec{f}(\vec{x})$ or, equivalently, according to the definition of the pseudoinverse,

$$J_{\vec{f}}^T(\vec{x})J_{\vec{f}}(\vec{x})\delta\vec{x} = -J_{\vec{f}}^T(\vec{x})\vec{f}(\vec{x}).$$

• This is Newtown where the Hessians H_{f_i} are taken equal to zero. This is justified when the residuals are small. Nearly quadratic convergence close to a solution.

Levenberg-Marquardt.

• Take the increment $\delta\vec{x}$ to be the solution of

$$[J_{\vec{f}}^T(\vec{x})J_{\vec{f}}(\vec{x}) + \mu\mathrm{Id}]\delta\vec{x} = -J_{\vec{f}}^T(\vec{x})\vec{f}(\vec{x}),$$

where the parameter μ is allowed to vary at each iteration.

• This is Newton where the term involving the Hessians is this time approximated by a multiple of the identity matrix. The Levenberg–Marqardt algorithm has convergence properties comparable to its Gauss–Newton cousin, but it is more robust.

Real Algebraic Problems

Example 1: What are the real roots of $x^4 - 3x^3 + x^2 - 5x + 1$?

Example 2: Do the surfaces defined in $\mathbb{R}^3$ by $x^2 - 5xy + 3z^3 = 0$ and $-2x^3 + y^3 + 2xyz - 1 = 0$ intersect?

Example 3: If the intersection is not empty, how does it look?

Example 4: When is the ellipse defined by

$$\frac{(x - x_0)^2}{a^2} + \frac{(y - y_0)^2}{b^2} - 1 = 0$$

inside the unit circle centered at the origin?

Examples 1 and 3 are *display* problems, and example 2 is a *query* problem, and example 4 is a *constraint* problem.

Query and constraint problems can be reduced to *quantifier elimination* problems *over the reals*:

- Ex. 2: $(\exists x)(\exists y)(\exists z)[x^2 - 5xy + 3z^3 = 0 \text{ and } - 2x^3 + y^3 + 2xyz - 1 = 0]$.
- Ex. 4: $(\forall x)(\forall y)[b^2(x - x_0)^2 + a^2(y - y_0)^2 - a^2b^2 = 0 \Longrightarrow x^2 + y^2 - 1 \leq 0]$.

Display problems may reduce to root finding but may also boil down to determining the topology of a set of varieties. Quantifier elimination plays a role there too.

Deflation

Idea: When a root α of $P(X)$ is found, we factor $P(X)$ into

$$P(X) = (X - \alpha)Q(X),$$

where $\deg Q(X) = \deg P(X) - 1$. The roots of $Q(X)$ are exactly the remaining roots of $P(X)$.

• A simple recurrence relation for deflation is easily found. Given

$$P(X) = a_n X^n + a_{n-1}X^{n-1} + \ldots + a_1 X + a_0,$$

and

$$Q(X) = b_{n-1}X^{n-1} + \ldots + b_0,$$

we clearly have:

$$\begin{cases} b_{n-1} = a_n, \\ b_{n-2} = a_{n-1} + \alpha b_{n-1}, \\ \ldots \\ b_0 = a_1 + \alpha b_1. \end{cases}$$

• In floating-point arithmetic, deflation is not an exact process, hence error creeps in and accumulates, making the computation of successive roots of a polynomial less and less accurate.

• It is a good idea to treat the roots of the successively deflated polynomials as *tentative* roots of the original polynomial $P(X)$. They are *polished* by taking them as (good) initial guesses for the Newton method applied to the original *nondeflated* polynomial.

Laguerre's Method

Idea: Design an iterative method guaranteed to converge to a root from any working point. Then use deflation to factor out this root, and iterate.

- Let $P(X)$ be a polynomial of degree n, define

$$G(X) = \frac{P'}{P}(X) \quad \text{and} \quad H(X) = [\frac{P'}{P}(X)]^2 - \frac{P''}{P}(X).$$

- Set x to some initial value x_0 and iterate $x \leftarrow x - a$ where

$$a = \frac{n}{G(x) + s\sqrt{(n-1)(nH(x) - G^2(x))}},$$

until a is small enough. Here $s = \mp 1$ is chosen at each iteration to maximize the magnitude of the denominator.

- **Note:** The computation of the radical requires complex arithmetic even when the coefficients of $P(X)$ are real.

Why Does It Work (Intuition)?

• Suppose that $P(X)$ factorizes over $\mathbb{C}$ as follows:

$$P(X) = (X - \alpha_1) \dots (X - \alpha_n).$$

It follows that

$$\frac{P'}{P}(X) = G(X) = \frac{1}{X - \alpha_1} + \dots + \frac{1}{X - \alpha_n},$$

and

$$[\frac{P'}{P}(X)]^2 - \frac{P''}{P}(X) = -G'(X) = H(X) = \frac{1}{(X - \alpha_1)^2} + \dots + \frac{1}{(X - \alpha_n)^2}.$$

• Now assume *(even though it is not true)* that the root α_1 that we seek is located at some distance a from our current guess x, while all other roots are assumed to be located at a distance b, i.e.,

$$a = x - \alpha_1, \quad b = x - \alpha_i, \quad \text{for } i = 2, \dots, n.$$

• This allows us to rewrite our constraints as

$$\begin{cases} \dfrac{1}{a} + \dfrac{n-1}{b} = G(x), \\ \dfrac{1}{a^2} + \dfrac{n-1}{b^2} = H(x). \end{cases}$$

• Solving this equation for a yields the Laguerre formulas.

Euclid's Algorithm

Theorem: Let k be a commutative field, and let $F(X)$ and $G(X)$ be two nonzero polynomials in one variable over k. Then there exist unique polynomials $Q(X), R(X)$ in $k[X]$ such that $F(X) = G(X)Q(X) + R(X)$ and deg $R(X) <$ deg $G(X)$.

Proof: Write

$$F(X) = a_n X^n + \ldots + a_1 X + a_0, \quad G(X) = b_d X^d + \ldots + b_1 X + b_0,$$

where $n =$ deg $F(X)$ and $d =$ deg $G(X)$ so that $a_n \neq 0$ and $b_d \neq 0$. We use induction on n.

If $n = 0$ and deg $G(X) >$ deg $F(X)$, we let $Q(X) = 0$ and $R(X) = F(X)$. If deg $F(X) =$ deg $G(X) = 0$, then we let $R(X) = 0$ and $Q(X) = a_n b_d^{-1}$.

Now assume the theorem proved for $n - 1$. We may assume deg $G(X) \leq$ deg $F(X)$ (otherwise take $Q(X) = 0$ and $R(X) = F(X)$). Then we can write

$$F(X) = a_n b_d^{-1} X^{n-d} G(X) + F_1(X)$$

where $F_1(X)$ has a degree strictly smaller than n. By the induction hypothesis, we can find $Q_1(X), R(X)$ such that

$$F(X) = a_n b_d^{-1} X^{n-d} G(X) + Q_1(X)G(X) + R(X)$$

and deg $R(X) <$ deg $G(X)$. Then we let

$$Q(X) = a_n b_d^{-1} X^{n-d} + Q_1(X)$$

which concludes the proof of existence of Q, R.

For uniqueness, suppose that

$$F(X) = Q_1(X)G(X) + R_1(X) = Q_2(X)G(X) + R_2(X)$$

with deg $R_1(X) <$ deg $G(X)$ and deg $R_2(X) <$ deg $G(X)$. Subtracting yields

$$(Q_1(X) - Q_2(X))G(X) = R_2(X) - R_1(X),$$

so that

$$\deg\,(Q_1(X) - Q_2(X)) + \deg\, G(X) = \deg\,(R_2(X) - R_1(X)).$$

Since deg $(R_2(X) - R_1(X)) <$ deg $G(X)$, this relation can only hold if $Q_1(X) - Q_2(X) = 0$, i.e., $Q_1(X) = Q_2(X)$ which in turns implies that $R_1(X) = R_2(X)$. QED.

Squarefree Polynomials

• Consider a polynomial $P(X)$ of degree n and its derivative $P'(X)$. Assuming that α is a root of multiplicity $m \leq n$ of $P(X)$, this polynomial can be written as

$$P(X) = (X - \alpha)^m Q(X),$$

where $Q(X)$ is a polynomial of degree $n - m$. Hence the derivative of $P(X)$ is

$$P'(X) = (X - \alpha)^{m-1}[mQ(X) + (X - \alpha)Q'(X)],$$

from which we deduce that α is a root of multiplicity $m - 1$ of P', and that $(X - \alpha)^{m-1}$ is a factor of the *greatest common divisor* (GCD) of $P(X)$ and $P'(X)$.

• This leads to the following definition: The *squarefree part* $P^*(X)$ of a polynomial $P(X)$ is the quotient of $P(X)$ by the greatest common divisor of $P(X)$ and its derivative $P'(X)$.

• The polynomials $P(X)$ and $P^*(X)$ have exactly the same roots, but all the roots of $P^*(X)$ are simple. If $P(X) = P^*(X)$, i.e., $P(X)$ has simple roots only, then $P(X)$ is said to be *squarefree*.

The Euclidean Algorithm

• Consider two integers x_0 and x_1, with $x_0 \geq x_1$. We can define the following sequence of divisions:

$$\begin{aligned} x_0 \quad &= \; a_1 x_1 + x_2, \\ x_1 \quad &= \; a_2 x_2 + x_3, \\ \ldots \\ x_{n-1} &= \; a_n x_n + 0. \end{aligned}$$

The sequence stops when the last remainder is zero, at which point x_n is the GCD.

• Example: Computing the GCD of $x_0 = 42$ and $x_1 = 30$. We have:

$$\begin{aligned} 40 &= 1 \times 30 + 12, \\ 30 &= 2 \times 12 + 6, \\ 12 &= 2 \times 6 + 0, \end{aligned}$$

and the GCD of 42 and 30 is 6.

• The same algorithm can be used to computed the GCD of two polynomials with rational coefficients.

Exact Representation of Algebraic Numbers

• When a real α is a root of some polynomial $P(X)$ with rational coefficients, it is said to be an *algebraic number*. Some real numbers are not algebraic; e.g., e, or π are *transcendental*.

• How can one represent *exactly* irrational algebraic numbers, such a $\sqrt{2}$?

• An *isolating interval* $[a, b]$ for α, i.e., an interval bounded by rational numbers $a < b$ that does not contain any other root of $P(X)$, together with the squarefree part $P^*(X)$ of $P(X)$, provides an exact representation of α.

• This representation also provides an arbitrarily precise numerical representation of α through bisection.

Sturm Sequences

• Let $P(X)$ be a squarefree polynomial and $P'(X)$ denote its derivative. The *Sturm sequence* of $P(X)$ is the sequence $\{F_i(X)\}$ of polynomials defined by

$$\begin{cases} F_0(X) = P(X), \\ F_1(X) = P'(X), \\ \text{for } i > 1, -F_i \text{ is the remainder obtained by dividing } F_{i-2} \text{ by } F_{i-1}. \end{cases}$$

Note: The degree of the polynomials F_i is strictly decreasing, ensuring that the sequence terminates in a finite number of steps with a constant polynomial F_k. This constant is different from zero since $P(X)$ is squarefree.

• **Theorem:** For any interval $[a, b]$ the number of roots of $P(X)$ in $[a, b]$ is $S(a) - S(b)$, where $S(x)$ is the number of sign changes in the sequence $[F_0(x), F_1(x), .., F_k(x)]$.

Note: When a and b are rational numbers, we can compute the number of roots in $[a, b]$ *exactly.*

• **Theorem:** Let $P(X)$ be a polynomial of degree n with rational coefficients, the number of real zeros of $P(X)$ that lie in any interval $[a, b]$ with rational bounds can be found in $O(n)$ arithmetic operations, after preprocessing that requires $O(n \log^2 n)$ arithmetic operations.

Root Isolation

• Given a squarefree polynomial $P(X)$ with rational coefficients, first find an interval $[a, b]$ that contains all its roots, e.g.,

$$b = -a = \max\{1 + \frac{p_i}{p_n} : i = 0, .., n-1\},$$

where p_i is the coefficient of the ith power of X in $P(X)$.

• Now, let $N = S(a) - S(b)$ denote the number of real roots of $P(X)$ in $[a, b]$. Use a binary search of $I = [a, b]$ to find a point c in I which separates it into two subintervals, each containing at least one of the roots of $P(X)$. This can be done in $O(n \log n)$ bisection steps, at each of which we must evaluate S, so that c can be found in $O(n^2 \log n)$ time.

• Apply the same process to each of the subintervals $[a, c]$ and $[c, b]$ to find isolating intervals for all the roots of $P(X)$. At most $N \leq n$ intervals will be processed, for a total cost of $O(n^3 \log n)$ (which dominates the cost of the Sturm sequence computation).

Example

• Consider the polynomial

$$P(X) = (X-1)X(X+1) = X^3 - X,$$

whose roots are $-1, 0, 1$.

• Its Sturm sequence is

$$\begin{cases} F_0(X) = P(X) = X^3 - X, \\ F_1(X) = P'(X) = 3X^2 - 1, \\ F_2(X) = 2/3X, \\ F_3(X) = 1. \end{cases}$$

Indeed, we have:

$$\begin{cases} F_0(X) = \dfrac{1}{3}XF_1(X) - F_2(X), \\ F_1(X) = \dfrac{9}{2}XF_2(X) - F_3(X). \end{cases}$$

• An interval $[a, b]$ containing all roots is given by

$$b = -a = \max\{1 + \frac{0}{1}, 1 + \frac{-1}{1}, 1 + \frac{0}{1}, 1 + \frac{1}{1}\},$$

yielding the interval $[-2, 2]$.

• Next we compute $S(-2)$ and $S(2)$. We have

$$\begin{cases} [F_0(-2) = -6, F_1(-2) = 11, F_2(-2) = -4/3, F_3(-2) = 1] \Longrightarrow S(-2) = 3, \\ [F_0(+2) = +6, F_1(+2) = 11, F_2(+2) = +4/3, F_3(+2) = 1] \Longrightarrow S(2) = 0, \end{cases}$$

which yields the (correct) number of roots in the interval $[-2, 2]$, $S(-2) - S(2) = 3$.

• After that, the interval $[-2, 2]$ can be subdivided until the three roots $-1, 0, 1$ are isolated.

Elimination Theory

Theorem: A necessary and sufficient condition for the system of n homogeneous linear equations in n unknowns

$$\begin{cases} a_{11}X_1 + \ldots + a_{1n}X_n = 0, \\ \ldots \\ a_{n1}X_1 + \ldots + a_{nn}X_n = 0. \end{cases}$$

to admit a non-trivial solution is that

$$\mathrm{Det}\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ & \ldots & \\ a_{n1} & \ldots & a_{nn} \end{pmatrix} = 0.$$

Theorem: A necessary and sufficient condition for the n homogeneous *polynomials* in $\mathbb{Q}[X_1, \ldots, X_n]$

$$\begin{cases} P_1(X_1, \ldots, X_n) = 0, \\ \ldots \\ P_n(X_1, \ldots, X_n) = 0, \end{cases}$$

to admit a common root is that their *Macaulay resultant* $R = \mathrm{Det}(A)$ be equal to zero, where A is a matrix whose entries are polynomials in the coefficients of the polynomials $P_i(X)$.

Sylvester Resultants

• Consider two polynomials $P(X)$ and $Q(X)$ in $\mathbb{Q}[X]$ defined by

$$\begin{cases} P(X) = a_n X^n + \ldots + a_1 X + a_0, \\ Q(X) = b_m X^m + \ldots + b_1 X + b_0. \end{cases}$$

• Multiply $P(X)$ successively by X^{m-1}, X^{m-2}, .. , X, and 1, and multiply $Q(X)$ by X^{n-1}, X^{n-2}, .. , X, 1. This yields the following linear system in the power products X^{n+m-1}, X^{n+m-2}, .., X, 1:

$$\begin{pmatrix} a_n & a_{n-1} & \ldots & a_1 & a_0 & & & & \\ & a_n & a_{n-1} & \ldots & a_1 & a_0 & & & \\ \ldots & & & & & & & & \\ & & & a_n & a_{n-1} & \ldots & a_1 & a_0 & \\ b_m & b_{m-1} & \ldots & b_1 & b_0 & & & & \\ & b_m & b_{m-1} & \ldots & b_1 & b_0 & & & \\ \ldots & & & & & & & & \\ & & & b_m & b_{m-1} & \ldots & b_1 & b_0 & \end{pmatrix} \begin{pmatrix} X^{n+m-1} \\ \ldots \\ X \\ 1 \end{pmatrix} = 0.$$

• Now if α is a common root of the polynomials $P(X)$ and $Q(X)$, then this system of homogeneous linear equations has a non-trivial solution, which implies in turn that the determinant of the $(m+n) \times (m+n)$ matrix formed by the coefficients of $P(X)$ and $Q(X)$ is zero. This is the Sylvester resultant of these two polynomials.

• This also works for polynomials with coefficients in $\mathbb{Q}[X_1, \ldots, X_p] = \mathbb{Q}[X_1, \ldots, X_{p-1}][X_p]$.

Bezout Resultants

• Consider two polynomials P_1, P_2 of degree n in X, and the determinant:

$$D(X,Y) = \begin{vmatrix} P_1(X) & P_2(X) \\ P_1(Y) & P_2(Y) \end{vmatrix}.$$

• This determinant vanishes whenever X is a common root of P_1, P_2 (the first row vanishes) and whenever $X = Y$ (the two rows are identical). It follows that the polynomial D is divisible by $(X - Y)$, and that the polynomial:

$$F(X,Y) = D(X,Y)/(X - Y)$$

vanishes whenever X is a common root of P_1, P_2. Clearly, F has degree $n-1$ in X and Y. We can rewrite F as a polynomial in Y, so that:

$$F(X,Y) = \sum_{i=0}^{n-1} F_i(X)Y^i,$$

where F_i is a polynomial of degree $n-1$ in X.

Bezout Resultants II

• Consider a common root α of $P_1(X)$ and $P_2(X)$. For any value Y, we have $F(\alpha, Y) = 0$, and it follows that all $F_i(\alpha)$ are equal to zero. Let F_i^j denote the coefficient of degree j of $F_i(X)$, we obtain the following linear system:

$$\begin{pmatrix} F_0(\alpha) \\ \vdots \\ F_i(\alpha) \\ \vdots \\ F_{n-1}(\alpha) \end{pmatrix} = M \begin{pmatrix} \alpha^0 \\ \vdots \\ \alpha^j \\ \vdots \\ \alpha^{n-1} \end{pmatrix} = 0,$$

where:

$$M = \begin{pmatrix} F_0^0 & \dots & F_0^j & \dots & F_0^{n-1} \\ \vdots & & \vdots & & \vdots \\ F_i^0 & \dots & F_i^j & \dots & F_i^{n-1} \\ \vdots & & \vdots & & \vdots \\ F_{n-1}^0 & \dots & F_{n-1}^j & \dots & F_{n-1}^{n-1} \end{pmatrix}.$$

• Considering the successive powers α^j as so many independent variables, it follows that this homogeneous linear system admits a non-trivial solution if and only if its determinant vanishes, i.e. $\mathrm{Det}(M) = 0$. This determinant is Bezout's resultant.

• This generalizes easily into a method for eliminating two variables among three polynomials (*Dixon resultants*).

Dixon Resultants

• Consider three polynomials P_1, P_2, P_3 in two variables X_1, X_2, with highest degree n in X_1 and m in X_2, and the determinant:

$$D(X_1, X_2, Y_1, Y_2) = \begin{vmatrix} P_1(X_1, X_2) & P_2(X_1, X_2) & P_3(X_1, X_2) \\ P_1(Y_1, X_2) & P_2(Y_1, X_2) & P_3(Y_1, X_2) \\ P_1(Y_1, Y_2) & P_2(Y_1, Y_2) & P_3(Y_1, Y_2) \end{vmatrix}.$$

• This determinant vanishes whenever (X_1, X_2) is a common root of P_1, P_2, P_3 (the first row vanishes), and also whenever $X_1 = Y_1$ (the first two rows are identical) or $X_2 = Y_2$ (the last two rows are identical). It follows that the polynomial D is divisible by $(X_1 - Y_1)(X_2 - Y_2)$, and that the polynomial:

$$F(X_1, X_2, Y_1, Y_2) = D(X_1, X_2, Y_1, Y_2)/((X_1 - X_2)(Y_1 - Y_2))$$

vanishes whenever (X_1, X_2) is a common root of P_1, P_2, P_3. Clearly, F has degree $n-1$ in X_1, $2m-1$ in X_2, $2n-1$ in Y_1, and $m-1$ in Y_2.

• We can rewrite F as a polynomial in Y_1 and Y_2, so that:

$$F(X_1, X_2, Y_1, Y_2) = \sum_{\substack{i=0 \\ j=0}}^{\substack{2n-1 \\ m-1}} F_{i,j}(X_1, X_2) Y_1^i Y_2^j,$$

where $F_{i,j}$ is a polynomial in X_1 and X_2, with degree in X_1 (resp. X_2) less than or equal to $n-1$ (resp. $2m-1$).

Dixon Resultants II

• Consider a common root (α_1, α_2) of P_1, P_2, P_3. For any value Y_1, Y_2, we have $F(\alpha_1, \alpha_2, Y_1, Y_2) = 0$, and it follows that all $F_{i,j}(\alpha_1, \alpha_2)$ are equal to zero. Let $F_{i,j}^{k,l}$ denote the coefficient of degree k in X_1 and l in X_2 of $F_{i,j}$, we obtain the following linear system:

$$\begin{pmatrix} F_{0,0}(\alpha_1, \alpha_2) \\ \vdots \\ F_{i,j}(\alpha_1, \alpha_2) \\ \vdots \\ F_{2n-1,m-1}(\alpha_1, \alpha_2) \end{pmatrix} = M \begin{pmatrix} \alpha_1^0 \alpha_2^0 \\ \vdots \\ \alpha_1^k \alpha_2^l \\ \vdots \\ \alpha_1^{n-1} \alpha_2^{2m-1} \end{pmatrix} = 0,$$

where

$$M = \begin{pmatrix} F_{0,0}^{0,0} & \dots & F_{0,0}^{k,l} & \dots & F_{0,0}^{n-1,2m-1} \\ \vdots & & \vdots & & \vdots \\ F_{i,j}^{0,0} & \dots & F_{i,j}^{k,l} & \dots & F_{i,j}^{n-1,2m-1} \\ \vdots & & \vdots & & \vdots \\ F_{2n-1,m-1}^{0,0} & \dots & F_{2n-1,m-1}^{k,l} & \dots & F_{2n-1,m-1}^{n-1,2m-1} \end{pmatrix}.$$

• Considering the successive powers $\alpha_1^k \alpha_2^l$ as so many independent variables, it follows that this homogeneous linear system admits a non-trivial solution if and only if its determinant vanishes, i.e., $\mathrm{Det}(M) = 0$. This determinant is Dixon's resultant.