Applications of Numerical Algebraic Geometry in Pure and Applied Mathematics

Chris Peterson - July 3, 2006

General Problem 1: Develop numeric tools for analyzing objects of interest to a pure mathematician.

General Problem 2: Develop algebraic and geometric tools for analyzing numerical data.

Two Natural reactions:

1) Data is too noisy for algebraic and geometric methods?

2) How can numerical methods work for algebraic objects. They are unstable under perturbations.

Example - Secant Varieties:

Let $I_A = k[A_1, A_2, \ldots, A_n]$ be the ideal of a variety V. The ideal, P, of the secant variety of V can be found by:

Algorithm:

Duplicate I_A to make an ideal, I_B, in the ring $k[B_1, B_2, \ldots, B_n]$

Form the ideal J whose generators are $\{C_i - (A_i s + B_i t)\}_{i=1\ldots n}$

Form the ideal $L = I_A + I_B + J$ in the ring $k[s, t, \{A_i, B_i, C_i\}_{i=1\ldots n}]$

Compute $P = L \cap k[C_1, C_2, \ldots, C_n]$.

Two remarks:

1) Higher secant varieties can be found by making more duplications.

2) This is computationally very expensive and thus impractical for all but the simplest objects!

Suppose we are only interested in the dimension of the r-secant variety. For instance, we would like to know if a certain secant variety is defective.

We could, of course, compute the ideal of the r-secant variety and then determine the degree of its associated Hilbert Polynomial.

Another way to compute the dimension of a variety is to compute the dimension of its tangent space at a smooth point of the variety.

Let V be a variety.

Let $P_1, P_2, \ldots, P_k$ be general points on V.

Let q be a general point in the span of these k points.

Then q is a general point on the k-secant variety, $\sigma_k(V)$, of V.

Terracini's Lemma: The tangent space to S_k at q is equal to the span of the tangent spaces to V at $P_1, P_2, \ldots, P_k$.

If we have the ideal I of V then we can compute a basis for the Tangent space to V at a point P by evaluating the Jacobian, $Jac(I)$ of I at P then taking the orthogonal complement.

I.e. $T_{V,P} = (Jac(I)_P)^{\perp}$.

Thus we can use Terracini's lemma if given $P_1, P_2, \ldots, P_k$ by putting all the T_{V,P_i} into one big matrix and computing its rank.

If the matrix corresponding to the span of the tangent spaces has a rank different than expected, then the secant variety is defective.

Terracini's Lemma is absolutely wonderful for certain objects like Segre, Grassmann and Veronese varieties.

Example: The Veronese surface in $\mathbb{P}^5$ has equations

$$e^2 - df, ce - bf, cd - be, c^2 - af, bc - ae, b^2 - ad$$

Evaluating the orthogonal complement of the Jacobian matrix at the points

$$[1 : 2 : 3 : 4 : 6 : 9] \text{ and } [1 : 3 : 5 : 9 : 15 : 25]$$

and placing the results into one big matrix yields a 6×6 matrix whose rank is 5.

This suggests (but does not prove) that the Veronese has a degenerate secant variety.

In general, there are some problems with the approach given above.

Problem 1: Computing the rank at a true "general point" is not really possible.

"Solution": Don't really pick a general point, just pick a "general enough point".

Problem 2: Computing the rank of a large matrix at a "sufficiently general" point over $\mathbb{Q}$ is computationally very expensive due to coefficient blow up.

"Solution": Don't compute the rank over $\mathbb{Q}$, instead compute the rank over a field with "sufficiently high" characteristic.

In this setting, unexpected drops in rank provide evidence rather than proof. Still, the results point us in the right direction.

The most serious problem is perhaps the following:

Problem 3: What if you don't have any points on V and can't hope to produce any?

Solution: ????

One possible solution, try to apply Terracini's lemma to points that are very, very near to the variety.

An obvious but interesting fact:

Evaluating the Jacobian matrix at points which are almost on a variety yields a matrix which is very close to the matrix obtained by evaluating the Jacobian matrix at a point exactly on the variety.

For instance, if your variety has codimension 3 and you evaluate the Jacobian matrix at a smooth point of the variety then your resulting matrix will have rank 3.

If you evaluate the Jacobian matrix at a point very, very close to a smooth point then your resulting matrix will have a "numerical rank" of 3.

In other words, only the first 3 singular values are numerically nonzero.

SINGULAR VALUE DECOMPOSITIONS

Given a matrix, $A \in \mathbb{C}^{m\times n}$, the singular value decomposition of A is a factorization

$A = U\Sigma V^*$ where

$U \in \mathbb{C}^{m\times m}$ is unitary

$\Sigma \in \mathbb{R}^{m\times n}$ is diagonal

$V \in \mathbb{C}^{n\times n}$ is unitary

The entries of Σ are non-negative real numbers arranged in decreasing order. The diagonal entries of Σ are called the singular values of A.

Theorem: Every matrix $A \in \mathbb{C}^{m\times n}$ has a singular value decomposition and the singular values of A are uniquely determined.

Proposition: The singular values of A are the square roots of the eigenvalues of A^*A (or equivalently of AA^*).

Proposition: The singular value decomposition of A can be used to write A as the sum of rank one matrices where each partial sum captures as much energy of A as possible in the sense of the Frobenius norm.

We can use the SVD to determine the closest matrix of lower rank.

We can determine how far away this lower rank matrix is from A.

Example: The Veronese surface in $\mathbb{P}^5$ has equations

$$e^2 - df, ce - bf, cd - be, c^2 - af, bc - ae, b^2 - ad$$

Evaluating the jacobian matrix at the points

$[1 : 2 : 3 : 4 : 6 : 9] + \epsilon[1 : 3 : 5 : 2 : 3 : 1]$ and

$[1 : 3 : 5 : 9 : 15 : 25] + \epsilon[2 : 1 : 4 : 3 : 4 : 1]$

yields two matrices which are very near a rank three matrix.

We can determine these two nearby rank 3 matrices by SVD.

We can compute their orthogonal complements and place the results into one big matrix.

This yields a 6×6 matrix which is very near a rank five matrix (again by SVD).

This suggests (but does not prove) that the Veronese has a degenerate secant variety.

How do we find points close to a variety?

We can use the method of Homotopy Continuation to find such points if we have the ideal of the variety.

Using Homotopy continuation together with an approximate Terracini yields a method for detecting when a variety has a defective secant variety.

The Principles of Homotopy Continuation

Given a polynomial system P, homotopy methods exploit the structure of P to construct a start system, P_0 that has sufficiently many regular solutions.

P_0 is embedded in the homotopy

$$H(\vec{\mathrm{x}}, t) = \gamma(1 - t)P_0(\vec{\mathrm{x}}) + tP(\vec{\mathrm{x}})$$

were $\gamma \in \mathbb{C}$ is a random number.

As t moves from 0 to 1, numerical continuation methods trace the paths that originate at the solutions of the start system towards solutions of the target system.

There are definitely some points to worry about:

Problem 1: How can we tell the difference between "small" singular values and singular values which are "numerically zero"?

Solution: We can repeat the Homotopy Continuation algorithm at a higher precision and see if the small singular values shrink.

Question: Shouldn't I be a bit suspicious?

Answer: Sure, but you should also be suspicious of many symbolic computations.

We are collecting evidence. Experience suggests it is easier to prove things that are "known" to be true than things that are not!

Problem 2: Won't the solutions of P_0 migrate to special solutions of P?

Solution: Take a general linear section to isolate a general point.

Problem 3: Won't the paths potentially "bump into" singularities along the homotopy?

Solution: Homotopies are real paths. So even if there is a (complex) codimension 1 singular locus, with probability one, you will miss it.

Some advantages of the approach outlined above:

1) Singular values can be determined for VERY large matrices.

2) Many numeric algorithms grow polynomially with the number of variables while symbolic algorithms tend to grow faster.

3) Homotopy Continuation is Parallelizable!

4) We can carry out these computations with varieties which have no rational points.

While we are going down the slippery slope of numerics, we may as well see where it takes us:

Consider the ideal, I, of the Veronese surface in $\mathbb{P}^5$.

If we perturb each generator of the ideal slightly, can we still use Terracini's lemma?

(Note that the ideal now is $\mathfrak{m}$-primary!)

We are asking the following:

Consider an ideal, I', that is very close to I and points that are very close to satisfying the conditions of I'. Can we use Terracini to "determine" the dimension of the secant varieties to $V(I)$?

Interesting Fact:

Let V be a variety defined by an ideal I.

Let p be a point on V.

Perturb I to get a new ideal I' and perturb p to get a point p'.

The Jacobian matrix of I' evaluated at p' is very close to the Jacobian matrix of I evaluated at p.

$$Jac(I + \epsilon J)_{P+\delta Q} =$$

$$Jac(I)_P + \epsilon Jac(J)_P + Jac(I)_{\delta Q} + \epsilon Jac(J)_{\delta Q}$$

Consequence: We can use I' and p' to understand the Tangent Space to V at p thus we can use Terracini's Lemma in this setting.

Note that given I, we can produce p' (but not p) by Homotopy Continuation.

Natural Question: Given I' can we produce p' by Homotopy Continuation?

Answer: Yes!

Small perturbations of I lead to small perturbations of the output of the Homotopy Continuation algorithm.

Let's continue down this slippery path.

First of all, if we know "enough" points on a variety do we know the variety?

In some sense, yes.

So let's define a **data variety** to be a very large "data matrix" where each row of the matrix satisfies the conditions imposed by an ideal.

Question: Can we use Terracini's Lemma to compute the dimension of the k-secant variety of a data variety?

Answer: If we have enough points on the variety to give a "good" approximation of the tangent space then Yes!

Question: How about if we perturb the data matrix?

Answer: If we had enough points to estimate the tangent space before the perturbation then we still have this property after a small perturbation.

Example: Pick your favorite 10,000 points on $\mathbb{P}^2$ and use them to make 10,000 points on the Veronese surface in $\mathbb{P}^5$.

Now we have a 10,000 $\times$ 6 matrix, A, which is the highly sampled Veronese.

Pick another 10,000 $\times$ 6 matrix, B, and consider $C = A + \epsilon B$.

Now view each row of C as a point, pick two of the points, P, Q, and find other points from your data set that are close to P and Q.

Put all of the nearest neighbors of P into a matrix and you will find it has 3 small singular values.

Similarly for Q.

Find the closest rank 3 matrix to P and its neighbors and the closest rank 3 matrix to Q and its neighbors.

Use these two rank 3 matrices to form a 6×6 matrix, F.

F will have one unexpectedly small singular value indicating that it is very close to a rank 5 matrix.

This provides evidence that the Veronese surface has a degenerate secant variety.

Question: Can we compute scheme and sheaf theoretic information numerically?

In particular, can we compute hilbert functions, hilbert polynomials, free resolutions, sheaf cohomology, etc.

Answer: I think so but am not completely sure.

Certainly some of these things can be computed.

A motivating philosophy is that many of the operations that can be carried out in symbolic algebra can be mimicked in numerics.

However, in the numerical setting you have an extra tool that you do not have in the symbolic setting: You can produce "generic" points.

This opens up new possibilities:

1) Terracini's Lemma can be applied in a more general setting.

2) Perhaps you can embed a curve by the "almost line bundle" corresponding to a point almost on the curve.

3) Given a square matrix you can produce an "Eigenscheme". A primary decomposition gives the information needed for the Jordan Canonical Form of the matrix. As a consequence, the Jordan Canonical Form can be computed numerically.

EIGENSCHEMES

The vector space of $(n+1) \times (n+1)$ matrices modulo the identity matrix is isomorphic to the vector space of sections of $\mathcal{T}_{\mathbb{P}^n}$.

Let S_A denote the section corresponding to a matrix A.

The zero locus of S_A determines a scheme, V_A and an ideal I_A.

Each of the schemes associated to a primary decomposition of I_A are supported on a linear space.

These linear spaces correspond to the eigenspaces.

The non-reduced components correspond to generalized eigenvectors.

A is diagonalizable iff I_A is a radical ideal.

Non-reduced components are Jordan Blocks.

Now to complete our slide into the numerical abyss/garden of Eden, let's approach things from a different point of view.

We have already considered a variety as a huge data matrix consisting of a large number of sample points on the variety.

Furthermore, information can be extracted even if there is a small bit of noise in the data.

Suppose you collect data such as EEG data, EKG data, weather data, or the result of some other experiment.

Can these data matrices be viewed as very noisy varieties?

Can the numerical tools that are derived from algebraic geometry be applied to these situations?

Answer: Perhaps sometimes.

I will close with one example:

ILLUMINATION SPACES (Joint with Geometry of Data Group at CSU)

Fix an object and fix a position of a digital camera.

For each possible illumination of the object you can produce a digital image.

A digital image is a collection of pixels with numerical values attached to each pixel.

We can string out the pixels and view the image as a point in a high dimensional vector space.

As we vary the illumination we obtain various points in this high dimensional space.

Pick an object.

If 1) The object is convex.

2) The object is illuminated from a distance.

3) The object is “Lambertian”.

Then

a) The first few spherical harmonics capture most of the energy of the data.

b) The set of all possible illuminations forms a convex set.

This has the consequence that the set of digital images lie very close to a low dimensional vector space.

A face is a reasonable approximation of such an object.

Through experimentation, we have found that the digital pictures of a fixed person in a fixed pose but under varying illuminations lie very close to a 10 dimensional linear space.

So each person under these restrictions corresponds to a point on GR(10,N) where N is large.

There are various metrics we can use to measure distances between points on the Grassmannian.

Some metrics that we used were successful in allowing us to differentiate between 50 different people.

Summarizing Statements:

1) Geometric and Algebraic methods can be of use to a mathematician interested in Applications.

But .....

2) Numerical methods can also be of use to a mathematician in Geometry and Algebra.

3) There is ample room for research. Many of the present tools and results in these hybrid areas are “stone age” level!