Hilbert's Nullstellensatz *

Graeme Taylor

May 23, 2005

Varieties and Ideals

For a ground field k, an affine algebraic variety is, roughly speaking, a set of points (essentially, N-tuples) $Y \subset k^N$ (I'll denote this A^N from now on) such that there exists a finite collection of polynomials $F_1 \dots F_r$ drawn from the polynomial ring $k[X_1, \dots, X_N]$ such that a point $x = (x_1, \dots, x_N)$ is in Y if and only if $F_1(x) = \dots = F_r(x) = 0$. This seems like a useful construct for a 'zero places theorem'!

In other words, we seek to describe geometric objects (regions in affine space) in terms of polynomials, to which we can then apply the full force of algebra. But an interesting problem involves the opposite process- given a collection of polynomials, where, if anywhere, do they all vanish?

It turns out that a good way to approach this is in terms of ideals, since if a collection of points evaluate at zero on some finite set of polynomials, then they'll also give a zero for the most likely infinite assortment of ways of combining those polynomials via addition, multiplication from the ring, and so on. There are hence two processes at work:

> **Definition**: Given an affine algebraic variety $Y \subset A^N$, the ideal $I(Y) \subset k[X_1, \dots, X_N]$ is given by $F \in I$ iff $F(x) = 0 \ \forall x$.
> **Definition**: Given an ideal I of $k[X_1, \dots, X_N]$, the corresponding variety $V(I)$ is the set $\{x \in A^N | F(x) = 0 \ \forall F \in I\}$.
> Note that I initially defined a variety in terms of a finite set of polynomials, to make the conversion from geometry to algebra more palatable. Fortunately, this is still the case here- whatever awkward and messily infinite ideal I you might offer, by Hilbert's basis theorem the polynomial ring is Noetherian and thus I will be finitely generated. That in turn ensures the existence of finitely many polynomials $F_1, \dots, F_r$ such that $I =< F_1 \dots F_r >$ and thus $V(I) = \{x \in A^N | F_1(x) = \dots = F_r(x) = 0\}$.

*First appeared on Everything2, at *http://www.everything2.com/index.pl?node_id=1724650*

Why do we need a Nullstellensatz?

The above definitions look similar almost to the point of being circular. So, one might reasonably expect that $I(V(J)) = J$ or $V(I(Y)) = Y$; that is, these processes are mutual inverses. Sadly, that isn't quite true. Considering the set-theoretic notion of an inverse image, this isn't too surprising- properties such as injectivity and surjectivity guide us as to when images and inverses behave nicely when layered. The following examples show that moving between ideals and varieties (formulae and points) we can gain objects (more formulae that evaluate to zero, more points that disappear on the set of polynomials).

Example 1: $Y \neq V(I(Y))$

For simplicity, take the ground field to be $\mathbb{C}$ the complex numbers and restrict our attention to a single dimension ($N = 1$). Then the integers $\mathbb{Z}$ are a perfectly good subset of $\mathbb{C}$; so we'll let them be our variety Y.

So $I(Y)$ is the set of polynomials that vanish for all integers- that is, every integer is a root of such polynomials. Sounds like a tall order? There's only one such 'polynomial', the zero polynomial. So $I(Y) = \{0\}$.

You may have spotted the problem now. If not, consider $V(I(Y))$. This, by the above, is $V(\{0\})$- the set of points in $\mathbb{C}$ such that, when evaluated at that point, the zero polynomial gives zero. That's all of them, i.e., $V(I(\mathbb{Z})) = \mathbb{C}$.

In general therefore $Y \neq V(I(Y))$, although $Y \subset V(I(Y))$.

Example 2: $J \neq I(V(J))$

Sticking with $k = \mathbb{C}$ and $N = 1$, we consider the ideal J generated by the polynomial X^2. That is, $J =< X^2 >$. So $V(J)$ is all the complex numbers c such that $c^2 = 0$; there's only one of those, 0. So $V(J) = \{0\}$.

Now we look at $I(V(J)) = I(\{0\})$. That's the collection of polynomials which evaluate to zero at zero. Certainly J is contained in that set; but there are more- any linear polynomial of the form aX will also satisfy this condition. So now $I(V(J)) =< X > \neq < X^2 >$.

Again we can only conclude $J \subset I(V(J))$, not equality.

These simple examples aren't quite pathological- there was no particular reason I needed the set $\{0\}$ in each case, it's just the most striking case. Factors that do matter are those such as algebraic closure, nilpotent elements, and reducibility. So having hopefully convinced you of the need for a Nullstellensatz, I'd best supply one.

Hilbert's Nullstellensatz for varieties and ideals

If k is algebraically closed and I is an ideal of $k[X_1, \ldots, X_N]$ then:

- $V(I) = \emptyset \Rightarrow I = k[X_1, \ldots, X_N]$
- $I(V(I))$ is the radical ideal of I, the set of all f in $k[X_1, \ldots, X_N]$ such that some power of f is in I.

Proof

For the first part we argue by contraposition- that any ideal I which isn't the whole ring generates a non-empty variety $V(I)$. So let I be such an ideal. The polynomial ring $k[X_1, \ldots, X_N]$ is Noetherian, so there is some maximal ideal M containing I (possibly I itself, but we don't care). Thus, there is a point $a = (a_1, \ldots, a_N)$ from A^N such that $M =< X_1 - a_1, \ldots, X_N - a_N >$, that is, $M = \{f | f(a) = 0\}$. So, a is in $V(M)$. But any point which vanishes on a set M must vanish on any subset of M, such as I- that is, a is in $V(I)$. Hence, $V(I)$ is non-empty and the first part of the theorem holds.

Now for the second part, pick some I and ideal of $k[X_1, \ldots, X_N]$ and an f from $I(V(I))$ (forgive the notation!).

We construct a larger ring in $N + 1$ variables, $k[X_1, \ldots, X_N, Y]$, and take the ideal $J =< I, 1 - Yf(x) >$. Suppose there is some $(a_1, \ldots, a_N, b)$ in $V(J)$. Then $(a_1, \ldots, a_N)$ is in $V(I)$ and $1 - bf(a_1, \ldots, a_N) = 0$ by construction. Since f is from $I(V(I))$, it vanishes on $V(I)$ and in particular at $(a_1, \ldots, a_N)$ so $1 = 0$. This is clearly absurd in any ring, so there cannot be such a point $(a_1, \ldots, a_N, b)$. In other words, we have shown $V(J) = \emptyset$ and thus, by the first part, $J = k[X_1, \ldots, X_N, Y]$.

In particular, therefore, 1 is in J. So we can chose some q_i from $k[X_1, \ldots, X_N]$ and q, p_i from $k[X_1, \ldots, X_N, Y]$ so that $1 = (\Sigma p_i q_i) + q(1 - Yf)$, where the q_i are generators for I.

We introduced Y as an arbitrary variable; now we fix Y to be $1/f(x)$ to obtain $1 = \Sigma p_i(x, 1/f(x))q_i(x)$

Set $r = max\{deg_Y\{p_i(x, Y)\}\}$ and multiply each side by $f(x)^r$ to clear denominators: $f(x)^r = \Sigma f(x)^r p_i(x, 1/f(x))q_i(x)$. Observe that the RHS is an element of I, so the LHS is. That is, f is in the radical of I.

Finally, suppose f is in the radical of I. Then f^r is in I and hence in $I(V(I))$. But a polynomial ring of a field is an integral domain, so if $f^r(p) = 0$, then $f(p) = 0$, i.e. f is in $I(V(I))$. Hence, $I(V(I)) =$ the radical of I and the proof is complete.