Evaluation ideal and variety for a trio of error independentbinaryclassifiers

Introduction

This notebook will detail the algebraic geometry computations that take us from the “evaluation ideal” created from the voting patterns of a trio of binary classifiers to the “evaluation variety”. An evaluation ideal is a set of polynomials connecting observable voting pattern frequencies by the classifiers to unknown sample statistics of the ground truth that are our evaluation goal. We want to “grade” the classifiers using only the frequencies of their voting patterns.

That “grade” exists in sample statistics space. The test has already been taken. We have the decisions of the judges. We are faced with the task of grading them now. Not in the future, not in the past. This is another example of how the task of evaluation is much simpler than that of training. We have to estimate something that already exists, if you will. And there is only one time we have to do it. Training is much harder. You must create judges that, in the future, will behave correctly. And they have to do it many times. The task of evaluation is trivial in comparison. Why have we not conquered this much simpler space of the whole enterprise of learning?

Algebraic geometry of three error independent binary classifiers

The mathematics of algebraic evaluation is algebraic geometry. Every algebraic evaluation problem can be stated as a polynomial system relating observable decision events to unknown sample statistics. Here we are going to define that polynomial system assuming that the classifiers made errors independently on the sample. This is “the spherical cow” of Evaluation Land - the simplifying assumption that allows you to proceed forward and carry out computations that give you insight into the original problem. Workers in Training Land also have a preferred spherical cow - “consider an identically, independently drawn sample”. It may take some getting used to this new cow if you are a new visitor from Training Land.

The evaluation ideal of three error independent binary classifiers

Algebraic geometry is mainly the study of the connection between sets of polynomials and geometric objects in the variable space of those polynomials. The sets of polynomials are called “polynomial ideals". A set of linear equations is also a polynomial ideal. We define the "evaluation ideal" of our evaluation to be,

In[3165]:=

```
Clear[MakeIndependentVotingIdeal]
MakeIndependentVotingIdeal[{i_, j_, k_}] :=
```

$$\{P_\alpha P_{i,\alpha} P_{j,\alpha} P_{k,\alpha} + (1 - P_\alpha)(1 - P_{i,\beta})(1 - P_{j,\beta})(1 - P_{k,\beta}) - f_{\alpha,\alpha,\alpha},$$
$$P_\alpha P_{i,\alpha} P_{j,\alpha}(1 - P_{k,\alpha}) + (1 - P_\alpha)(1 - P_{i,\beta})(1 - P_{j,\beta}) P_{k,\beta} - f_{\alpha,\alpha,\beta},$$
$$P_\alpha P_{i,\alpha}(1 - P_{j,\alpha}) P_{k,\alpha} + (1 - P_\alpha)(1 - P_{i,\beta}) P_{j,\beta}(1 - P_{k,\beta}) - f_{\alpha,\beta,\alpha},$$
$$P_\alpha P_{i,\alpha}(1 - P_{j,\alpha})(1 - P_{k,\alpha}) + (1 - P_\alpha)(1 - P_{i,\beta}) P_{j,\beta} P_{k,\beta} - f_{\alpha,\beta,\beta},$$
$$P_\alpha(1 - P_{i,\alpha}) P_{j,\alpha} P_{k,\alpha} + (1 - P_\alpha) P_{i,\beta}(1 - P_{j,\beta})(1 - P_{k,\beta}) - f_{\beta,\alpha,\alpha},$$
$$P_\alpha(1 - P_{i,\alpha}) P_{j,\alpha}(1 - P_{k,\alpha}) + (1 - P_\alpha) P_{i,\beta}(1 - P_{j,\beta}) P_{k,\beta} - f_{\beta,\alpha,\beta},$$
$$P_\alpha(1 - P_{i,\alpha})(1 - P_{j,\alpha}) P_{k,\alpha} + (1 - P_\alpha) P_{i,\beta} P_{j,\beta}(1 - P_{k,\beta}) - f_{\beta,\beta,\alpha},$$
$$P_\alpha(1 - P_{i,\alpha})(1 - P_{j,\alpha})(1 - P_{k,\alpha}) + (1 - P_\alpha) P_{i,\beta} P_{j,\beta} P_{k,\beta} - f_{\beta,\beta,\beta}\}$$

One convention in algebraic geometry may bother you. Ultimately we are interested in the geometrical object these polynomials define in the finite space needed for evaluating three independent binary classifiers. We want to consider the points in sample statistics space where all these equations are zero. We are really interested in these equations,

In[3167]:=

```
MakeIndependentVotingIdeal[{1, 2, 3}] // Map[(# == 0) &, #] &
```

Out[3167]=

$$\{P_\alpha P_{1,\alpha} P_{2,\alpha} P_{3,\alpha} + (1 - P_\alpha)(1 - P_{1,\beta})(1 - P_{2,\beta})(1 - P_{3,\beta}) - f_{\alpha,\alpha,\alpha} == 0,$$
$$P_\alpha P_{1,\alpha} P_{2,\alpha}(1 - P_{3,\alpha}) + (1 - P_\alpha)(1 - P_{1,\beta})(1 - P_{2,\beta}) P_{3,\beta} - f_{\alpha,\alpha,\beta} == 0,$$
$$P_\alpha P_{1,\alpha}(1 - P_{2,\alpha}) P_{3,\alpha} + (1 - P_\alpha)(1 - P_{1,\beta}) P_{2,\beta}(1 - P_{3,\beta}) - f_{\alpha,\beta,\alpha} == 0,$$
$$P_\alpha P_{1,\alpha}(1 - P_{2,\alpha})(1 - P_{3,\alpha}) + (1 - P_\alpha)(1 - P_{1,\beta}) P_{2,\beta} P_{3,\beta} - f_{\alpha,\beta,\beta} == 0,$$
$$P_\alpha(1 - P_{1,\alpha}) P_{2,\alpha} P_{3,\alpha} + (1 - P_\alpha) P_{1,\beta}(1 - P_{2,\beta})(1 - P_{3,\beta}) - f_{\beta,\alpha,\alpha} == 0,$$
$$P_\alpha(1 - P_{1,\alpha}) P_{2,\alpha}(1 - P_{3,\alpha}) + (1 - P_\alpha) P_{1,\beta}(1 - P_{2,\beta}) P_{3,\beta} - f_{\beta,\alpha,\beta} == 0,$$
$$P_\alpha(1 - P_{1,\alpha})(1 - P_{2,\alpha}) P_{3,\alpha} + (1 - P_\alpha) P_{1,\beta} P_{2,\beta}(1 - P_{3,\beta}) - f_{\beta,\beta,\alpha} == 0,$$
$$P_\alpha(1 - P_{1,\alpha})(1 - P_{2,\alpha})(1 - P_{3,\alpha}) + (1 - P_\alpha) P_{1,\beta} P_{2,\beta} P_{3,\beta} - f_{\beta,\beta,\beta} == 0\}$$

But it is a pain to carry all these equal to zero notation around. So we drop it, and prefer to work with,

In[3168]:=

```
MakeIndependentVotingIdeal[{1, 2, 3}]
```

Out[3168]=

$$\{P_\alpha P_{1,\alpha} P_{2,\alpha} P_{3,\alpha} + (1 - P_\alpha)(1 - P_{1,\beta})(1 - P_{2,\beta})(1 - P_{3,\beta}) - f_{\alpha,\alpha,\alpha},$$
$$P_\alpha P_{1,\alpha} P_{2,\alpha}(1 - P_{3,\alpha}) + (1 - P_\alpha)(1 - P_{1,\beta})(1 - P_{2,\beta}) P_{3,\beta} - f_{\alpha,\alpha,\beta},$$
$$P_\alpha P_{1,\alpha}(1 - P_{2,\alpha}) P_{3,\alpha} + (1 - P_\alpha)(1 - P_{1,\beta}) P_{2,\beta}(1 - P_{3,\beta}) - f_{\alpha,\beta,\alpha},$$
$$P_\alpha P_{1,\alpha}(1 - P_{2,\alpha})(1 - P_{3,\alpha}) + (1 - P_\alpha)(1 - P_{1,\beta}) P_{2,\beta} P_{3,\beta} - f_{\alpha,\beta,\beta},$$
$$P_\alpha(1 - P_{1,\alpha}) P_{2,\alpha} P_{3,\alpha} + (1 - P_\alpha) P_{1,\beta}(1 - P_{2,\beta})(1 - P_{3,\beta}) - f_{\beta,\alpha,\alpha},$$
$$P_\alpha(1 - P_{1,\alpha}) P_{2,\alpha}(1 - P_{3,\alpha}) + (1 - P_\alpha) P_{1,\beta}(1 - P_{2,\beta}) P_{3,\beta} - f_{\beta,\alpha,\beta},$$
$$P_\alpha(1 - P_{1,\alpha})(1 - P_{2,\alpha}) P_{3,\alpha} + (1 - P_\alpha) P_{1,\beta} P_{2,\beta}(1 - P_{3,\beta}) - f_{\beta,\beta,\alpha},$$
$$P_\alpha(1 - P_{1,\alpha})(1 - P_{2,\alpha})(1 - P_{3,\alpha}) + (1 - P_\alpha) P_{1,\beta} P_{2,\beta} P_{3,\beta} - f_{\beta,\beta,\beta}\}$$

Dropping the notation has no effect. Algebraic manipulations of the above set (multiplying them together, etc.) would be equivalent to polynomials of zero for points that satisfied the input set.

One “forest for the trees” note: the above ideal is one many possible ones. Algebraic evaluation is very much like data streaming algorithms. You are creating a sketch of the decisions by an ensemble of noisy judges. In the case of binary classification considered here, that “data sketch” is the frequency of their item-by-item voting patterns. Other polynomial systems are possible even for independent binary classifiers. We could be trying to evaluate sample statistics that look at how judges evaluated two different sample items, for example.

The evaluation variety of three error independent binary classifiers