Detection of a Gravitropism Phenotype in glutamate receptor-like 3.3 Mutants of Arabidopsis thaliana Using Machine Vision and Computation

Nathan D. Miller, Tessa L. Durham Brooks, Amir H. Assadi and Edgar P. Spalding

Copyright © 2010 by the Genetics Society of America
DOI: 10.1534/genetics.104.118711
FIGURE S1.—Solution vectors by two different methods. The solution vector that minimizes Eq. 1 when applied to the data obtained in condition C is shown in black. A highly similar result obtained by an alternative method similar to that described by Fisher (1936) is shown in orange. The alternative method treats an entire time course as a linear combination of tip angle measurements (TA) at each time point multiplied by a unitless coefficient $\lambda$. A given trial of 301 time points can thus be represented by a single number $X$:

$$X = \lambda_1 \cdot TA(1) + \lambda_2 \cdot TA(2) + \ldots + \lambda_{301} \cdot TA(301)$$

An equivalent statement is that $X$ is the dot product of the $\lambda$ and TA vectors, $X = \lambda \cdot TA$. The design of the experiments performed here allowed each genotype sampled with $n$ trials to be characterized by a population of $nX$ values. In Fisher’s discriminant analysis and the version used to create the above orange line, the set of $\lambda$ values that maximizes the difference between the mutant and wild type distributions of $X$ values were found. The square of the student's t-statistic was the function Fisher and we used to determine the distance between the $X$ distributions being discriminated. When the derivative of this function with respect to $\lambda$ is 0, the distance between the distributions is maximum. A least squares method was used to find the set of $\lambda$ values, the vector $\lambda$, that nullified the derivative of the squared t-test function. The resulting $\lambda$ values that best separates the mutant from the wild type (orange line above).

Fisher (1936) sought a linear combination of sepal length and width measurements that best separated the species *Iris setosa* from *Iris versicolor*. Our application required finding a linear combination of tip angle measurements that best separated two groups from each other (wild type and *glr3.3*) with the added constraint that the resulting did not separate the two *glr3.3* alleles from each other. This added constraint, separating A from B/C instead of A from B, was the reason we chose the minimax optimizer over the Fisher LDA method as the primary approach described in the text. But Fisher's original solution described here produced the equivalent result (orange line in the above figure) when the data from the two mutant alleles were pooled and a mutant (A) from wild type (B) discrimination was performed. The resulting solution vector did not separate the two mutant alleles from each other when they were subjected to an A from B discrimination (data not shown). Before the discrimination procedures, PCA was used to reduce the dimensionality of the data from 301 down to the 6 dimensions that contained 99% of the variance. The result was projected back into 301 dimensions before plotting in the figure.
FIGURE S2.—Solution vectors for separating mutant variances or means from wild type. The vector of weighting coefficients that separated the mutant and wild-type means (black line) was similar in shape to that which found the mutant variances to be larger than the wild type (purple line). However, they were not functionally interchangeable. The mean-splitting vector could not achieve the variance separation of Eq. 2, nor could the variance splitting vector achieve the mean separation of Eq. 1.