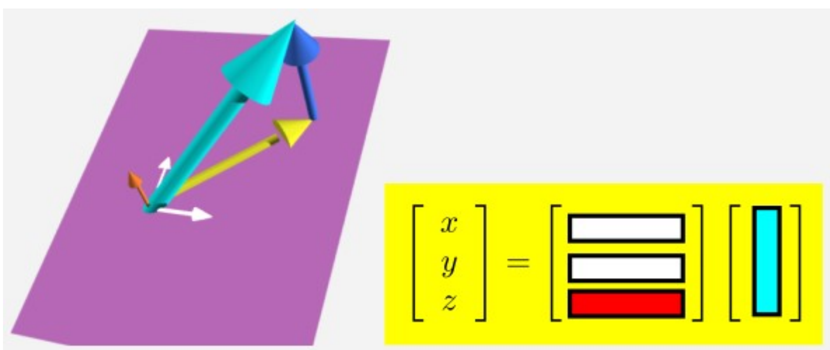
Chi-square and Projection

Arunsoumya Basu

Roll No. bs2029

We return to our familiar framework, the projection of a vector of, say, $\mathbb{R}^3$ onto a subspace of dimension, say, 2. But this time we will work with the probability distributions of the lengths of the random vector and its projections.



For convenience, we denote vectors by the initials of their colours, so the picture contains $\vec{c}$ (cyan), its projection $\vec{y}$ (yellow) on the violet subspace V, the perpendicular $\vec{b}$ dropped from $\vec{c}$ to V (blue), the orthonormal bases $\vec{w}_1, \vec{w}_2$ of V and the orthonormal basis $\vec{r}$ (red) of $V^{\perp}$.

Suppose in the above figure, $\vec{c}$ is distributed as $N_3(\vec{0}, \mathbf{I_3})$, i.e., its components are i.i.d. $N(0, 1)$ random variables. So its squared norm is distributed as χ^2_3, being the sum of squares of 3 i.i.d. standard normal random variables. We now want to find the distributions of the squared norms of $\vec{y}$ and $\vec{b}$.

For this, we first consider an orthogonal transformation $\vec{c} \rightsquigarrow \vec{u}$ brought about by an orthogonal matrix $\mathbf{A}$, i.e., $\underbrace{\begin{bmatrix} x \\ y \\ z \end{bmatrix}}_{\vec{u}} = \underbrace{\begin{bmatrix} \vec{w_1} \\ \vec{w_2} \\ \vec{r} \end{bmatrix}}_{\mathbf{A}} \vec{c}.$

Then, $\vec{u} \sim N_3(\vec{0}, \mathbf{I_3})$, so its components are i.i.d. standard normal random variables. Now, we note that $\vec{y} = \vec{w_1} \cdot \underbrace{\langle \vec{w_1}, \vec{c} \rangle}_{x} + \vec{w_2} \cdot \underbrace{\langle \vec{w_2}, \vec{c} \rangle}_{y}$, so

$||\vec{y}||^2 = ||xw_1||^2 + ||yw_2||^2 = x^2 \cdot ||w_1||^2 + y^2 \cdot ||w_2||^2 = (x^2 + y^2) \sim \chi^2_2$, as $x, y \overset{i.i.d.}{\sim} N(0, 1)$. [Here y (which has been used to denote a real number) and $\vec{y}$ (which has been used to denote a vector) are not to be confused.]

Similarly, $\vec{b} = \vec{r} \cdot \underbrace{\langle \vec{r}, \vec{c} \rangle}_{z}$, so $||\vec{b}||^2 = z^2 \sim \chi^2_1$. Moreover, $(x^2 + y^2)$ is independent of z^2, so the distributions of the squared norms of $\vec{y}$ and $\vec{b}$ are independent.

All these match our expectations: $\vec{y}$ lies in V, a subspace of dimension 2, so the corresponding degrees of freedom is also 2, and $\vec{b}$ lies in $V^\perp$ of dimension 1, so the corresponding degrees of freedom is 1. Further, the fact that $\vec{y}$ and $\vec{c}$ are orthogonal translates to the independence of the probability distributions of their squared norms.

In general, suppose $\vec{Y} \sim N_n(\vec{0}, \mathbf{I_n})$ and $\mathbb{R}^n = V_1 \oplus V_2 \oplus \cdots \oplus V_k$ ($\oplus$ denotes the direct sum of subspaces). Then $\vec{Y} = P_{V_1}\vec{Y} + P_{V_2}\vec{Y} + \cdots + P_{V_k}\vec{Y}$, where $P_{V_i}\vec{Y}$ is the projection of $\vec{Y}$ onto V_i for $i = 1, 2, \cdots, k$. Then, we use an argument similar to the one used before (for $n = 3$ and $k = 2$): we start with orthonormal bases of each of $V_1, \cdots, V_k$, then use an orthogonal transform that converts $\vec{Y}$ to the vector of lengths of projections along the bases, and finally conclude that $||P_{V_i}\vec{Y}||^2 \sim \chi^2_{\dim(V_i))}$ $\forall\, i \in \{1, 2, \cdots, k\}$. Also, the distributions of these projection vectors are independent of one another, since V_i and V_j are orthogonal spaces for $i \neq j$.

A slightly more general version of this is when $\vec{Y} \sim N_n(\vec{0}, \sigma^2\mathbf{I_n})$. In that case, we can use the previous setup for $\dfrac{\vec{Y}}{\sigma}$ to arrive at the fact $||P_{V_i}\vec{Y}||^2 \sim \sigma^2\chi^2_{\dim(V_i))}\ \forall\ i \in \{1, 2, \cdots, k\}$, the constant σ^2 appearing in RHS as every length has been scaled up by a factor of σ, compared to the $N_n(\vec{0}, \mathbf{I_n})$ version.

In an attempt to obtain an even more generalised result, we can incorporate a non-zero mean vector, i.e., $\vec{Y} \sim N_n(\vec{\mu}, \sigma^2\mathbf{I_n})$. We will then obtain the result $||P_{V_i}\vec{Y}||^2 \sim \sigma^2\chi^2_{\dim(V_i))}(||P_{V_i}\vec{\mu}||^2)\ \forall\ i \in \{1, 2, \cdots, k\}$, i.e., a non-centrality parameter given by $||P_{V_i}\vec{\mu}||^2$, which is the squared norm of the mean projected on V_i.

This has an interesting consequence: if we project $\vec{Y}$ onto a subspace to which $\vec{\mu}$ is orthogonal, the squared norm of the projection will be a central chi-square random variable, the non-centrality parameter being 0. This is a very useful result that will be used for testing of hypothesis in the upcoming sections.