HW Addition: Mean and Variance of a Projection

We will suppose we have p points in $\mathbb{R}^2$, although once we do the projection, they all become scalars. We will store the points in a $p \times 2$ matrix A.

If we project the data to the $x-$axis, the mean and variance is the mean and variance of the data in the first column of A.

If we project the data to the $y-$axis, the mean and variance is the mean and variance of the data in the second column of A.

Now suppose we project the data to the vector $\mathbf{u}$. You may assume that $\|\mathbf{u}\| = 1$ so that the formula to project one point $\mathbf{x}$ would be:

$$\text{Proj}_{\mathbf{u}}(\mathbf{x}) = (\mathbf{x} \cdot \mathbf{u})\ \mathbf{u}$$

We will look at the mean and variance of the *scalar projection*, $\mathbf{x} \cdot \mathbf{u}$.

For example, suppose we have three points:

x	1	2	−1
y	0	1	1

$$\mathbf{u} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Then the new scalar values from the projection are:

$$\frac{1+0}{\sqrt{2}}, \qquad \frac{2+1}{\sqrt{2}}, \qquad \frac{-1+1}{\sqrt{2}}$$

from which we can compute a mean and variance. Note that in matrix form, we could have written:

$$A\mathbf{u} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix}$$

We want to investigate the relationship of the mean and variance to the projection. To do that, let's set our notation to be the following:

$$A = \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ \vdots & \vdots \\ x_p & y_p \end{bmatrix} \qquad \mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \qquad \text{Assume } \|\mathbf{u}\| = 1$$

Let the mean of the first column of A be denoted by $\bar{x}$, and mean of the second by $\bar{y}$.

Here are some questions to answer:

1. Show that the mean of the projection is the projection of the mean.

2. In linear algebra, if $\mathbf{b}$ is a column vector, then what operation(s) should I perform in order to sum the squared elements of $\mathbf{b}$? (Write it out)

3. If the first and second columns of A have been mean-subtracted, what is the matrix-vector form of the variance of the projection? Your answer to the previous question may be helpful.

4. Show that you're correct by writing a short script in Maple using 100 randomly placed points in the plane, and the same vector $\mathbf{u}$ as in our example above. That is, compute the mean and variance of the projection by first performing the projection, then show that your answers correspond to the computations in exercises 1 and 3 above.