HW 1: Matlab with histograms on Fisher's data

(Due Sep 5)

Grading Notes: 30 points total with 5 pts each.

I'm looking to see that you were able to get Matlab to work, and that you were able to see some histograms of the data. In particular,

- Use the `size` command to see how many flowers.
- Use the `hist` command to plot a histogram.
- Compute mean and variance.
- Mean-subtract the data, divide by the standard deviation.
- Scatter plot $(1,2)$, then $(1,3)$ then $(1,4)$. Use different colors for different flowers, and different symbols.
- Check out linear discriminant and confusion matrix.

HW 2: Statistics calculations, p 20 of Day 1 notes

(Due Sep 7, 30 pts- 5 pts each except 6)

5. We should find that $\bar{x} = -0.1$ and $\bar{y} = 0.5$, and the covariance is 0.84.

6. (Also done in class)

$$\bar{x} = \frac{1}{p}\sum_{i=1}^{p} x_i \quad \Rightarrow \quad \frac{1}{p}\sum_{i=1}^{p}(ax_i) = a\frac{1}{p}\sum_{i=1}^{p} x_i = a\bar{x} \qquad \frac{1}{p}\sum_{i=1}^{p}(x_i + b) = \bar{x} + b$$

   Therefore, if $\bar{x}$ is the mean of data in vector $\mathbf{x}$, then $a\bar{x} + b$ is the mean of $a\mathbf{x} + b$.

7. Since we define variance as the following, we can use the previous answer to find the mean of the different sets.

$$\sigma^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

   The variance of $a\mathbf{x} + b$ is given by:

$$\frac{1}{n-1}\sum_{i=1}^{n}((ax_i + b) - (a\bar{x} + b))^2 = a^2\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = a^2\sigma^2$$

   This formula works for all scalars a and b.

8. Similar formulae for the covariance:

$$\mathrm{Cov}(ax,y) = \frac{1}{n-1}\sum_{i=1}^{n}(ax_i - a\bar{x})(y_i-\bar{y}) = a\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y}) = a\mathrm{Cov}(x,y)$$

$$\mathrm{Cov}(x,by) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})(by_i - b\bar{y}) = b\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y}) = b\mathrm{Cov}(x,y)$$

9. $\mathrm{Cov}(x,a) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})(a-a) = 0$

10. Because the mean of $x+a$ is $\bar{x}+a$, the a's will cancel in covariance, as does b:

$$\mathrm{Cov}(x+a,y+b) = \frac{1}{n-1}\sum_{i=1}^{n}((x_i+a)-(\bar{x}+a))((y_i+b)-(\bar{y}+b)) =$$

$$\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y}) = \mathrm{Cov}(x,y)$$

11. (a) If X and Y are the same data, then the covariance becomes the variance, and the correlation coefficient simplifies to 1.

    (b) If $Y = mX$, then the covariance was computed earlier, and was $m^2S_x^2$. The correlation coefficient would be:

$$r_{xy} = \frac{S_{xy}^2}{S_xS_y} = \frac{m^2S_x^2}{|m|S_x|m|S_x} = \frac{m^2}{|m|^2}$$

    which we interpret as 1 if $m > 0$, -1 if $m < 0$, and if $m = 0$, the correlation is undefined. The exact same thing happens if $Y = mX + b$.

    NOTE: Where does the absolute value of x come in? The standard deviation is the square root of the variance, so

$$S_{mx} = \sqrt{S_{mx}^2} = \sqrt{m^2S_x^2} = |m|S_x$$

HW 3: Matlab, Hanford data with Linear Regression

(Due Sep 10- 10 points total.)

SOLUTION: We should form the normal equations and solve. The slope was approximately 9.2 and the intercept was approximately 114.68. Also: Be sure the plot looks nice!

HW 4: Matlab and the n-armed bandit - Run the Matlab code.

(Due Sep 13- 5 points total)

HW 5: Variance and Projections (Some Matlab, some by hand)

(Due Sep 14 - 15 points total)

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

   Let matrix A be $p \times 2$ and hold the data. Let $\mathbf{u}$ be a unit vector onto which we perform the projection. Then the projected scalar data will be given by the $p \times 1$ vector

   $$A\mathbf{u} = \begin{bmatrix} x_1u_1 + y_1u_2 \\ x_2u_1 + y_2u_2 \\ x_3u_1 + y_3u_2 \\ \vdots \\ x_pu_1 + y_pu_2 \end{bmatrix}$$

   To see what the projected mean is, imagine summing down that column- We can write that as:

   $$\frac{1}{p}\left(u_1\sum_{i=1}^{p} x_i + u_2\sum_{i=1}^{p} y_i\right) = \left[\frac{1}{p}\sum_{i=1}^{p} x_i, \frac{1}{p}\sum_{i=1}^{p} y_i\right] \cdot \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

   The expression to the left is the mean of the projection, the expression to the right is the projection of the mean.

2. To get the sum of the squared elements, if $\mathbf{b}$ is a column vector, we take $\mathbf{b}^T\mathbf{b}$.

3. Using the previous expressions, the variance will be $\frac{1}{p-1}(A\mathbf{u})^T(A\mathbf{u}) = \frac{1}{p-1}(\mathbf{u}^TA^TA\mathbf{u})$

4. (Verification in Matlab)

HW 6: Exercises 1, 2, 5 in n-armed bandit (p 12)

(Due Sep 17 - Sorry about the notation! 15 points total)

1. If $P(1), P(2), P(3)$ are three probabilities (machine a, b, c), and $P(1)$ is max, show that the updated values still sum to 1.

   $P_{t+1}(1) = P_t(1)+\beta(1-P_t(1)) \qquad P_{t+1}(2) = P_t(2)+\beta(0-P_t(2)) \qquad P_{t+1}(3) = P_t(3)+\beta(0-P_t(3))$

   Summing these together, and noting the $P_t(1) + P_t(2) + P_t(3) = 1$, then we get:

   $$1 + \beta - \beta P_t(1) - \beta P_t(2) - \beta P_t(3) = 1 + \beta - \beta = 1$$

2. TYPO: $0 < \beta < 1$. Show the updated values stay between 0 and 1.

- Do the three numbers stay positive? Yes. For the winner, the probability is less than 1, so we're adding two positive numbers together. For the second and third machines, we can re-write the expression as

  $$P_{t+1} = (1 - \beta)P_t$$

  so that these are both positive as well.
- Are the numbers less than 1? In the first case, we add a slight amount that is proportional to the distance between P_1 and 1, so that sum will remain less than 1. In the second two cases, the new values can be written as a fraction of the previous case. If they began less than one, they are even smaller the second time around.

5. This should really only be done after 4. If you play around with the values a bit, we see that:

$$P_2 = (1 - \beta)P_1 + \beta$$
$$P_3 = (1 - \beta)P_2 + \beta = (1 + \beta)^2 P_1 + \beta(1 - \beta) + \beta$$
$$P_4 = (1 - \beta)^3 + \beta\left((1 - \beta)^2 + (1 - \beta) + 1\right)$$

and so on so that

$$P_{t+1} = (1 - \beta)^t + \beta \sum_{i=1}^{t-1} (1 - \beta)^t$$

As $t \to \infty$, $(1 - \beta)^t \to 0$ (if $0 \le \beta \le 1$), and the sum is a geometric series! The infinite sum simplifies to:

$$\beta \left(\frac{1}{1 - (-(1 - \beta))} \right) = 1$$

(Hopefully you were able to see some kind of pattern!)

HW 7: Functions v Scripts (all Matlab)

(Due Sep 19- 3 problems in Matlab. 15 points)

HW 8: Knapsack Problem in Matlab

(Due Sep 24- Solve the knapsack problem. 10 points)

HW 9: Linear Algebra Computations

(Due Sep 21 - 18 points total- 3 pts each)

1. Perform row reduction on the augmented matrix (augment with both x's) to get

$$\left[\begin{array}{cc|cc} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

   To interpret this, we see that column three (a pivot column) is linearly independent to columns 1 and 2 (the other pivot columns), so the corresponding vector is not in H. The last column is in H, and we see that $\mathbf{x}_2 = 2\mathbf{v}_1 + \mathbf{v}_2$.

2. The isomorphism is straightforward, but needs to be stated explicitly. The function is defined via the coordinate mapping:

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 \in \mathbb{R}^3 \quad \longleftrightarrow \quad (\alpha_1, \alpha_2) \in \mathbb{R}^2$$

3. For the projector, we need orthonormal columns, so first we take our matrix U as:

$$U = \begin{bmatrix} 1/\sqrt{14} & 3/\sqrt{10} \\ 3/\sqrt{14} & -1/\sqrt{10} \\ -2/\sqrt{14} & 0 \end{bmatrix}$$

   The projector is then UU^T, which unfortunately doesn't simplify much (you can use a calculator here):

$$UU^T \approx \begin{bmatrix} 0.97 & -0.08 & -0.14 \\ -0.08 & 0.74 & -0.43 \\ -0.14 & -0.43 & 0.29 \end{bmatrix}$$

   The projection onto the plane is $UU^T\mathbf{x}$, or about $(0.69, -0.94, 0.43)$. The distance between the point and the plane is the magnitude of the vector:

$$\|UU^T\mathbf{x} - \mathbf{x}\| \approx 1.86$$

7. The dot product should be straightforward to compute to show orthogonality:

$$(P\mathbf{x} - \mathbf{x}) \cdot \mathbf{u} = \left(\frac{\mathbf{x}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u} - \mathbf{x}\right) \cdot \mathbf{u} = 0$$

3, p 56 If Q has o.n. columns, then we'll recall that $Q^TQ = I$, and:

$$\|Q\mathbf{x}\|^2 = (Q\mathbf{x})^T(Q\mathbf{x}) = \mathbf{x}^TQ^TQ\mathbf{x} = \mathbf{x}^T\mathbf{x} = \|\mathbf{x}\|^2$$

4, p 56 Continue the example to get three vectors, and so on.