Scientific Computing for Biologists

Linear Algebra Review I

Instructor: Paul M. Magwene

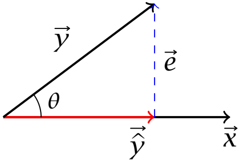
09 September 2014

Overview of Lecture

- ANOVA as bivariate regression
- Partial correlation
- Introduction to Matrices
  - Matrices as collections of vectors
  - Special matrices
- Matrix operations
  - Matrix addition, subtraction
  - Matrix multiplication
  - Transpose
  - More special matrices
- Matrices as linear transformations
- Linear dependence/independence
- Matrix inverses
- Solving simultaneous linear equations

Geometry of Bivariate Regression

Geometric interpretation of regression as projection:



$$\vec{\hat{y}} = b\vec{x}$$

$$b = \frac{\vec{x} \cdot \vec{y}}{\vec{x} \cdot \vec{x}} \tag{1}$$

Two-group ANOVA as Regression

We can also use a geometric perspective to test whether the mean of a variable differs between two groups of subjects.

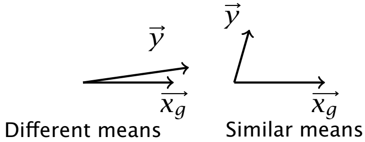
- Setup a 'dummy variable' as the predictor X_g. We assign all subjects in group 1 the value 1 and all subjects in group 2 the value -1 on the dummy variable. We then regress the variable of interest, Y, on X_g.

$$\mathsf{y} = \mathsf{X}_g \mathsf{b} + \mathsf{e}$$

	Raw		Centered	
Group	Y_i	X_i	y_i	x_i
1	2	-1	-3	$-\frac{4}{3}$
	3	-1	-1	$-\frac{4}{3}$
2	5	1	0	$\frac{2}{3}$
	6	1	1	$\frac{2}{3}$
	6	1	1	$\frac{2}{3}$
	7	1	2	$\frac{2}{3}$
Mean	5	$\frac{1}{3}$	0	0

Two-group ANOVA as Regression, cont

- When the means are different in the two groups, X_g will be a good predictor of the variable of interest, hence $\vec{y}$ and $\overrightarrow{x_g}$ will have a small angle between them.
- When the means in the two groups are similar, the dummy variable will not be a good predictor. Hence the angle between $\vec{y}$ and $\overrightarrow{x_g}$ will be large.



Geometry of the Population Regression Model

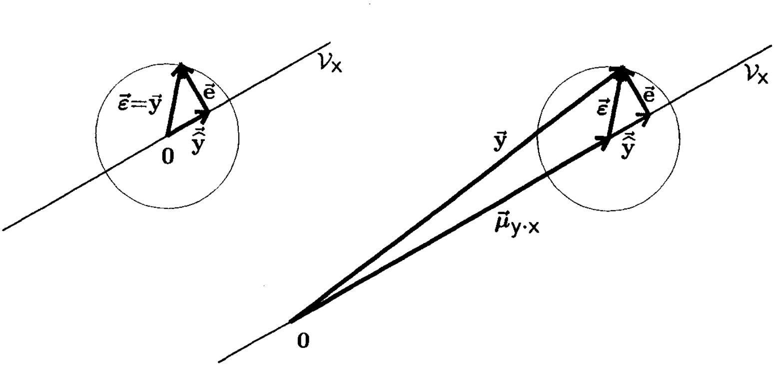


Figure: **Left:** null hypothesis of no regression effects is true; **Right:** null model is false.

Is my regression significant? $\implies$ Is $|\vec{\hat{y}}|^2$ large relative to $|\vec{e}|^2$?

Comparing the Effect Space and the Error Space

To compare the squared length of $|\vec{\hat{y}}|^2$ and $|\vec{e}|^2$ we divide them by the dimension of the subspaces in which they lie.

$$M(\vec{\hat{y}}) = \frac{|\vec{\hat{y}}|^2}{\dim(\mathcal{V}_x)}$$
$$M(\vec{e}) = \frac{|\vec{e}|^2}{\dim(\mathcal{V}_e)}$$

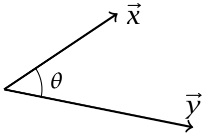
We compare these by defining a statistic, F:

$$F = \frac{M(\vec{\hat{y}})}{M(\vec{e})} = \frac{\dim(\mathcal{V}_e)|\vec{\hat{y}}|^2}{\dim(\mathcal{V}_x)|\vec{e}|^2} = \frac{(N-p-1)R^2}{p(1-R^2)}$$

When null hypothesis is true, $F \approx 1$; when it is false, $F \gg 1$.

Reminder: Correlation

Last time we saw that correlation is a measure of association and is a function of the angle between two vectors (variables).



$$\text{cor}(X, Y) = r_{XY} = \cos\theta = \frac{\vec{x} \cdot \vec{y}}{|\vec{x}||\vec{y}|}$$

Q: Correlation is a pairwise measure. Can we extend this idea to more than two variables?

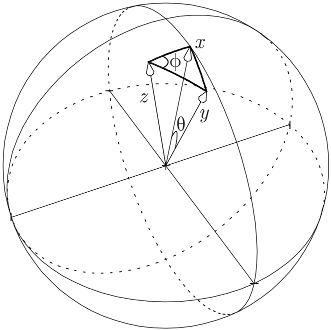
Partial Correlation

The partial correlation of $\vec{x}$ and $\vec{y}$ given $\vec{z}$ is the correlation of $\vec{x}$ and $\vec{y}$ after we “account for” their joint association with $\vec{z}$.

What we mean by “account for” here is essentially ‘to factor out’, by projecting $\vec{x}$ and $\vec{y}$ onto $\vec{z}$ and then calculating the correlation of the residual vectors.

Geometry of Partial Correlation

The partial correlation of $\vec{x}$ and $\vec{y}$ given $\vec{z}$ is equivalent to the correlation of the residuals after projecting $\vec{x}$ and $\vec{y}$ onto $\vec{z}$.



$$\mathrm{cor}(X, Y|Z) = r_{XY.Z} = \mathrm{cor}(\hat{x}_{\perp Z}, \hat{y}_{\perp Z}) = \cos\phi$$

Algebra of Partial correlation

Algebraicly, one can calculate the partial correlation between X and Y given Z as:

$$\text{cor}(X, Y|Z) = r_{XY.Z} = \frac{r_{XY} - r_{XZ}r_{YZ}}{\sqrt{(1 - r_{XZ}^2)(1 - r_{YZ}^2)}}$$

This extends logically when Z represents a set of variables rather than just a single variable.

$$\text{cor}(X, Y|Z, W) = r_{XY.ZW} = \frac{r_{XY.Z} - r_{XZ.W}r_{YZ.W}}{\sqrt{(1 - r_{XZ.W}^2)(1 - r_{YZ.W}^2)}}$$

Introduction to Matrices

- One way to think about a matrix is as a collection of vectors. This is, in essence, what a multivariate data set is.
- A matrix which has n rows and p columns will be referred to as a $n \times p$ matrix. $n \times p$ is the shape of the matrix.

$$\underset{(n \times p)}{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{bmatrix}$$

Special Matrices

- Zero matrix

$$0 = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

- Square matrix
  A matrix whose shape is is $n \times n$

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

- Ones matrix

$$1 = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}$$

- Diagonal matrix
  A square matrix where the off-diagonal elements are zero.

$$A = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

Scalar Multiplication of a Matrix

- Let k be a scalar and let A be the matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{bmatrix}$$

- then

$$kA = \begin{bmatrix} ka_{11} & ka_{12} & \cdots & ka_{1p} \\ ka_{21} & ka_{22} & \cdots & ka_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ ka_{n1} & ka_{n2} & \cdots & ka_{np} \end{bmatrix}$$

Addition and Subtraction of Matrices

- Let A and B be matrices that have the same shape, $n \times p$:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{bmatrix} \quad B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{bmatrix}$$

- then

$$A + B = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1p} + b_{1p} \\ a_{21} + b_{11} & a_{22} + b_{22} & \cdots & a_{2p} + b_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} + b_{n1} & a_{n2} + b_{n2} & \cdots & a_{np} + b_{np} \end{bmatrix}$$

$$A - B = A + (-B)$$

Multiplying a Matrix by a Vector

- Let A be a $n \times p$ matrix, and let x be a $p \times 1$ vector

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{bmatrix} \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}$$

- then

$$A\mathsf{x} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1p}x_p \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2p}x_p \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{np}x_p \end{bmatrix}$$

Note that $A\mathsf{x}$ is a vector with shape $n \times 1$. The i-the element of $A\mathsf{x}$ is equivalent to the dot product of the i-th row vector of A with x.

General Matrix Multiplication

- Let A be a $n \times p$ matrix and B be a $p \times q$ matrix:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{bmatrix} \quad B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1q} \\ b_{21} & b_{22} & \cdots & b_{2q} \\ \vdots & \vdots & \vdots & \vdots \\ b_{p1} & b_{n2} & \cdots & b_{nq} \end{bmatrix}$$

- The product AB is an $n \times q$ matrix whose (i, j)-entry is the dot product of the i-th row vector of A and the j-th column vector of B.

Matrix Arithmetic Rules

i. $A + B = B + A$
ii. $(A + B) + C = A + (B + C)$
iii. $k(A + B) = kA + kB$
iv. $(kA)B = k(AB)$
v. $(AB)C = A(BC)$ (associative)
vi. $A(B + C) = AB + AC$ (distributive)
vii. $(A + B)C = AC + BC$ (distributive)

Alert

*Matrix multiplication is **not** commutative, i.e. $AB \neq BA$ in general.*

Be careful when you expand expressions like $(A + B)(A + B)$.

Matrix Transpose

- We denote the transpose of a matrix as A^T
- If A is an $n \times p$ matrix, then A^T is a $p \times n$ matrix where $A^T_{ji} = A_{ij}$
- Transpose rules:
  - $(A^T)^T = A$
  - $(A+B)^T = A^T + B^T$
  - $(AB)^T = B^T A^T$

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{bmatrix}$$

$$A^T = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & \vdots & \vdots \\ a_{1p} & a_{12} & \cdots & a_{np} \end{bmatrix}$$

More Special Matrices

Symmetric matrix – square matrix, A, where $A^T = A$

Skew-symmetric matrix – square matrix, A, where $A^T = -A$

Identity Matrix – diagonal matrix, I, where

$$I = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

- $IA = AI = A$ if I and A are $n \times p$ matrices
- $A = I\mathbf{x}$ is a diagonal matrix where $a_{ii} = x_i$ if I is an $n \times n$ matrix and $\mathbf{x}$ is a $n \times 1$ vector.

Orthogonal matrix – square matrix for which $A^T A = AA^T = I$.

Matrices as Linear Transformations

- Let A be a particular $n \times p$ matrix. Than for any p-vector x, the product $A\mathsf{x}$ is a n-vector.
- We say that the matrix A determines a function from $\mathbb{R}^p$ to $\mathbb{R}^n$.
  - $A(k\mathsf{x}) = k(A\mathsf{x})$ where k is a scalar.
  - If y is also a p-vector than $A(\mathsf{x} + \mathsf{y}) = A\mathsf{x} + A\mathsf{y}$ is an n-vector
- A function, f, where $f(\mathsf{x} + \mathsf{y}) = f(\mathsf{x}) + f(\mathsf{y})$ and $f(k\mathsf{x}) = kf(\mathsf{x})$ is called a ***linear transformation***.

Highlight

Every matrix determines a linear transformation!

Every linear transformation can be represented by a matrix!

Examples of Linear Transformation in $\mathbb{R}^2$

- reflection in the x-axis

$$\begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} x \\ -y \end{bmatrix}$$

- reflection in the line $y = x$

$$\begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} y \\ x \end{bmatrix}$$

- shear parallel to the x-axis

$$\begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} x + ay \\ y \end{bmatrix}$$

- projection onto the x-axis

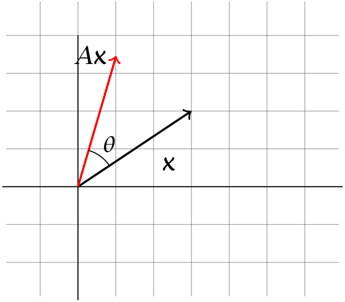
$$\begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} x \\ 0 \end{bmatrix}$$

- How about reflection in the y-axis? shear parallel to the y-axis? projection onto the y-axis?

Examples of Linear Transformation: Rotation

- The rotation of the plane, by an angle θ about the origin is given by:

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$



Linear dependence/independence

- You'll remember that a *linear combination* of vectors is an equation of the form $z = b_1 x_1 + b_2 x_2 + \cdots + b_p x_p$
- A list of vectors, $x_1, x_2, ..., x_p$, is said the be ***linearly dependent*** if there is a non-trivial combination of them which ie qual to the zero vector.

$$b_1 x_1 + b_2 x_2 + \cdots + b_p x_p = 0$$

- A list of vectors that are not linearly dependent are said to be ***linearly independent***

Matrix Inverses

- If A is a *square matrix* and C is a matrix of the same size where $AC = I$ and $CA = I$ than C is the inverse of A and we denote is A^{-1}.

$$AA^{-1} = A^{-1}A = I$$

- Rules for inverses:
  - Only square matrices are invertible
  - A matrix for which we can find an inverse is called ***invertible*** (non-singular)
  - A matrix for which no inverse exists is ***singular*** (non-invertible)
  - If A and B are both invertible $p \times p$ matrices than $AB^{-1} = B^{-1}A^{-1}$ (note change in order).

Highlight

If a matrix is invertible than it's columns form a linearly independent list of vectors!

More facts about Matrix Inverses

- Not every square matrix is invertible
- Every orthogonal matrix is invertible
- Any diagonal matrix, A, where the a_{ii} are non-zero, is invertible

Simultaneous Linear Equations

- A set of simultaneous linear equations are equations like the following:

$$\begin{aligned} x_1 + 3x_2 + 2x_3 &= 3 \\ -x_1 + x_2 + 2x_3 &= -2 \\ 2x_1 + 4x_2 - 2x_3 &= 10 \end{aligned}$$

- Simultaneous linear equations have either:
  - No solutions
  - One solution
  - Infinitely many solutions

Matrices and Simultaneous Linear Equations

- Matrices can be used to represent and solve simultaneous linear equations. For example,

$$\begin{aligned} x_1 + 3x_2 + 2x_3 &= 3 \\ -x_1 + x_2 + 2x_3 &= -2 \\ 2x_1 + 4x_2 - 2x_3 &= 10 \end{aligned}$$

  Can be represented by the equation $A\mathbf{x} = \mathbf{h}$:

$$\begin{bmatrix} 1 & 3 & 2 \\ -1 & 1 & 2 \\ 2 & 4 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ -2 \\ 10 \end{bmatrix}$$

- Solve this equation by pre-multiplying both sides of the equation by A^{-1}.

$$\begin{aligned} A^{-1}A\mathbf{x} &= A^{-1}\mathbf{h} \\ \mathbf{x} &= A^{-1}\mathbf{h} \end{aligned}$$

Simultaneous Equations and Matrix Inverses

- $A\mathsf{x} = \mathsf{h}$ has a unique solution iff A is invertible.
- If A is a singular matrix than $A\mathsf{x} = \mathsf{h}$ either has no solution or infinitely many solutions.