Linear Algebra

Matrix Basics

Fully Reduced Matrix

The following matrix is in row-reduced echelon form:

$$\begin{bmatrix} \mathbf{1} & 2 & 0 & 1 & 0 & 2 & | & 2 \\ 0 & 0 & \mathbf{1} & 2 & 0 & 0 & | & 3 \\ 0 & 0 & 0 & 0 & \mathbf{1} & 1 & | & 4 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{0} & | & 0 \end{bmatrix}$$

Columns 1, 3 and 5 contain leading entries, and so their corresponding vectors form the basis for the span of this set of vectors. The final row has no leading entry, and so is parametized as shown below. If the unknowns being solved for are $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$ then:

$$\begin{aligned} \alpha_1 &= 2 - 2 - 1 - 2t \\ \alpha_2 &= 3 - 2 \\ \alpha_3 &= 4 - t \\ \alpha_4 &= t \end{aligned}$$

Adjacency Matrix

If A is an $n \times n$ adjacency matrix of a graph and $a_{i,j}^k$ represents the (i, j) entry of A^k, then $a_{i,j}^k$ is equal to the number of walks of length k from V_i to V_j.

Transpose

$$\begin{aligned} (A+B)^T &= A^T + B^T \\ (AB)^T &= B^T A^T \\ (cA)^T &= cA^T \\ \det(A^T) &= \det(A) \\ (A^T)^{-1} &= (A^{-1})^T \end{aligned}$$

Determinant

$$\begin{aligned} \det \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} &= (-1)^{1+3} \det[A_{13}] + (-1)^{2+3} \det[A_{23}] + (-1)^{3+3} \det[A_{33}] \\ &= (-1)^{1+3} \det \begin{bmatrix} d & e \\ g & h \end{bmatrix} + (-1)^{2+3} \det \begin{bmatrix} d & e \\ g & h \end{bmatrix} + (-1)^{3+3} \det \begin{bmatrix} d & e \\ g & h \end{bmatrix} \end{aligned}$$

$$\det(AB) = \det(A)\det(B)$$

$$\det(A^{-1}) = \frac{1}{\det(A)}$$

$$\det(A) = \det(A^T)$$

$$\det(cA) = c^n \det(A)$$

$$\det \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix} = -\det \begin{pmatrix} b_1 & b_2 \\ a_1 & a_2 \end{pmatrix}$$

$$\det \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix} = \det \begin{pmatrix} a_1 & a_2 \\ b_1 + ca_1 & b_2 + ca_2 \end{pmatrix}$$

Euclidean Vector Space

Projections

Projection of $\tilde{v}$ onto $\tilde{u} = (\tilde{v} \cdot \tilde{u})\tilde{u}$

Projection of $\tilde{v}$ onto the direction of $\tilde{u} = \frac{(\tilde{v} \cdot \tilde{u})\tilde{u}}{\|\tilde{u}\|}$

Areas

$$\text{Area parallelogram} = \|\tilde{u} \times \tilde{v}\| = \|\tilde{u}\|\|\tilde{v}\| \sin\theta \text{ (if in R}^3\text{) or } \left|\det\begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix}\right| \text{ (if in R}^2\text{)}$$

$$\text{Area parallelepiped} = \|\tilde{w} \cdot (\tilde{u} \times \tilde{v})\| = \left|\det\begin{pmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix}\right|$$

$$\text{Area of triangle with vertices } (\tilde{w}, \tilde{u}, \tilde{v}) = \frac{1}{2}\|(\tilde{u} - \tilde{w}) \times (\tilde{v} - \tilde{w})\|$$

Vectors

$$\|\tilde{u}\|\|\tilde{v}\| \cos\theta = \tilde{u} \cdot \tilde{v}$$
$$\tilde{u}, \tilde{v} \text{ orthogonal if } \theta = \pi/2, \cos\theta = 0, \tilde{u} \cdot \tilde{v} = 0$$

Lines

Parametric form: $x = a + bt, \; y = c + dt, \; z = e + ft$

Cartesian form: $\frac{x-a}{b} = \frac{y-d}{d} = \frac{z-e}{f}$

Vector form: $(x, y, z) = (a, c, e) + t(b, d, f)$

$$(x_1, y_1, z_1) = (a_1, b_1, c_1) + t(a_2, b_2, c_2)$$
$$(x_2, y_2, z_2) = (a_3, b_3, c_3) + s(a_4, b_4, c_4)$$

Parallel if $t = ws$ for some $w \in R$

Intersection with line or plane: sub one line in and solve

Note: if given two points define one as the origin and convert the other into a vector.

Planes

Cartesian form: $ax + by + cz - d = 0 = (x, y, z) \cdot \tilde{n}$

Perpendicular vector: $\tilde{n} = (a, b, c)$

Vector form: $(x, y, z) = (x_0, y_0, z_0) + t(u_1, u_2, u_3) + s(v_1, v_2, v_3)$

When given a plane in Cartesian form and a point it passes through, vector form will be given by:

$$(x, y, z) = (x_0, y_0, z_0) + t(0,1,a) + s(1,0,b)$$

Where (x_0, y_0, z_0) is the given point, and a and b are found via $(0,1,a) \cdot \tilde{n} = (0,1,b) \cdot \tilde{n} = 0$.

General Vector Spaces

Subspaces

A subspace of R^n is a subset S of R^n such that:

- S is non-empty; $S \neq \emptyset$
- $\tilde{u}, \tilde{v} \in S \rightarrow (\tilde{u} + \tilde{v}) \in S$; closed under vector addition of inputs
- $\tilde{u} \in S, \alpha \in R \rightarrow \alpha\tilde{u} \in S$; closed under scalar multiplication of inputs

To prove that S is a subspace, all three properties must be proven to hold. To prove that it is not a subspace, only a single specific exception to one of the properties need be found. All spans are subspaces.

Span

The span is the set of all linear combinations of a given set of vectors.

$$\text{span}\{v_1, v_2, v_3\} = \{\alpha v_1 + \beta v_2 + \gamma v_3; \alpha, \beta, \gamma \in R\}$$

To determine if a particular vector $\tilde{v}_0$ is in a given span, solve the following linear system:

$$\tilde{v}_0 = \alpha_1 \tilde{v}_1 + \alpha_2 \tilde{v}_2 + \alpha_3 \tilde{v}_3$$

$$\begin{bmatrix} v_{1,1} & v_{2,1} & v_{3,1} & | & v_{0,1} \\ v_{1,2} & v_{2,2} & v_{3,2} & | & v_{0,2} \\ v_{1,3} & v_{2,3} & v_{3,3} & | & v_{0,3} \end{bmatrix}$$

An inconsistent system indicates that the vector does not belong to the span; otherwise it does.

$$\text{Dim(span)} = \text{rank}$$

Solution Space

The solution space is the set of all vectors that satisfy the equation:

$$A\tilde{v} = \tilde{0}, \text{where } A \text{ is the matrix of vectors in the system}$$

Note that the dimension of the solution set depends upon the number of vectors, not on the dimension of the vectors themselves.

To check if a given vector is in the solution space of a linear system, multiply it by the matrix forming that linear system A, and check to see that it equals zero.

$$\text{Dim(solution space)} = \text{nullity}$$

Basis

A basis is a set of linearly independent vectors that span a given subspace. In order words, a basis for a subspace is a set of vectors necessary and just sufficient to span that subspace.

A given subspace will generally have many possible bases, so to define which basis it is being written in, a subscript can be included:

$$[\tilde{u}]_B = \alpha_1 \tilde{v}_1 + \alpha_2 \tilde{v}_2 + \alpha_3 \tilde{v}_3, \text{where } \tilde{v}_t \text{ are the basis vectors}$$

Row and Column Space

The row space of a matrix is the span of the columns (i.e. the columns are interpreted as vectors). The column space of a matrix is the span of the rows (i.e. the rows are interpreted as vectors).

A basis for the row space can be found by computing the RREF of the matrix and reading off the non-zero rows directly as the basis. The basis of column space can be found from the RREF selecting the columns with leading entries, and reading off the corresponding vectors (as columns) of the original matrix.

Inner Product Spaces

Inner Products

The inner product is a generalization of the dot product. In a vector space, it is a way to multiply vectors together, with the result of this multiplication being a scalar. While the dot product follows a particular rule, the inner product can follow many different rules, so long as the rule satisfies the following properties:

- $\langle \tilde{u}, \tilde{v} \rangle = \langle \tilde{v}, \tilde{u} \rangle$; symmetric
- $\langle \alpha\tilde{u}, \tilde{v} \rangle = \alpha\langle \tilde{u}, \tilde{v} \rangle$; linear with respect to scalar multiplication
- $\langle \tilde{u} + \tilde{w}, \tilde{v} \rangle = \langle \tilde{u}, \tilde{v} \rangle + \langle \tilde{w}, \tilde{v} \rangle$; linear with respect to vector addition
- $\langle \tilde{u}, \tilde{u} \rangle > 0$ unless $\tilde{u} = 0$, where $\langle \tilde{u}, \tilde{u} \rangle = 0$

When the inner product is written in matrix form $\langle \tilde{x}, \tilde{y} \rangle = \tilde{x}^T A \tilde{y}$, these criteria simplify to:

- $A = A^T$: symmetric
- $\langle \tilde{x}, \tilde{x} \rangle > 0$ unless $\tilde{u} = 0$, where $\langle \tilde{u}, \tilde{u} \rangle = 0$

The inner product can be written in three basic forms:

- $\langle \tilde{u}, \tilde{v} \rangle$, where u and v are vectors
- $\int_0^1 p(x)q(x)\,dx$, where $p(x)$ and $q(x)$ are polynomials
- $\tilde{u}^T A \tilde{v}$, where A is an $n \times n$ matrix and the vectors u, v have dimension n

Orthogonality

Two vectors u and v are said to be orthogonal if $\langle \tilde{u}, \tilde{v} \rangle = 0$. A set of vectors is orthogonal if each vector in the set is orthogonal to every other vector in the set. A set of vectors is called orthonormal if it is orthogonal, and all vectors in the set have unit length.

In inner product spaces, length is defined as: $\|\tilde{v}\| = \sqrt{\langle \tilde{v}, \tilde{v} \rangle}$.

Note that if a matrix B is composed of orthogonal column vectors, then it $B^T = B^{-1}$.

Orthogonal Projections

The orthonormal projection $\tilde{p}$ of a vector $\tilde{v}$ onto subspace W with basis $\{\tilde{u}_1, \tilde{u}_2, \ldots, \tilde{u}_t\}$ is given by:

$$\tilde{p} = \langle \tilde{v}, \tilde{u}_1 \rangle \tilde{u}_1 + \langle \tilde{v}, \tilde{u}_2 \rangle \tilde{u}_2 + \cdots + \langle \tilde{v}, \tilde{u}_t \rangle \tilde{u}_t$$

Note that if $\tilde{v}$ is in vector space W with an *orthonormal* basis $\{\tilde{e}_1, \tilde{e}_2, \ldots, \tilde{e}_t\}$ then $\tilde{v}$ can be written as:

$$\tilde{v} = \langle \tilde{v}, \tilde{e}_1 \rangle \tilde{e}_1 + \langle \tilde{v}, \tilde{e}_2 \rangle \tilde{e}_2 + \cdots + \langle \tilde{v}, \tilde{e}_t \rangle \tilde{e}_t$$

If the basis $\{\tilde{e}_1, \tilde{e}_2, \ldots, \tilde{e}_t\}$ is *not orthonormal*, then $\tilde{v}$ can be written as:

$$\begin{aligned}
\tilde{v} &= \left\langle \tilde{v}, \frac{\tilde{e}_1}{\|\tilde{e}_1\|} \right\rangle \frac{\tilde{e}_1}{\|\tilde{e}_1\|} + \left\langle \tilde{v}, \frac{\tilde{e}_2}{\|\tilde{e}_2\|} \right\rangle \frac{\tilde{e}_2}{\|\tilde{e}_2\|} + \cdots + \left\langle \tilde{v}, \frac{\tilde{e}_t}{\|\tilde{e}_t\|} \right\rangle \frac{\tilde{e}_t}{\|\tilde{e}_t\|} \\
&= \frac{1}{\|\tilde{e}_1\|^2} \langle \tilde{v}, \tilde{e}_1 \rangle \tilde{e}_1 + \frac{1}{\|\tilde{e}_2\|^2} \langle \tilde{v}, \tilde{e}_2 \rangle \tilde{e}_2 + \cdots + \frac{1}{\|\tilde{e}_t\|^2} \langle \tilde{v}, \tilde{e}_t \rangle \tilde{e}_t \\
&= \frac{1}{\langle \tilde{e}_1, \tilde{e}_1 \rangle} \langle \tilde{v}, \tilde{e}_1 \rangle \tilde{e}_1 + \frac{1}{\langle \tilde{e}_2, \tilde{e}_2 \rangle} \langle \tilde{v}, \tilde{e}_2 \rangle \tilde{e}_2 + \cdots + \frac{1}{\langle \tilde{e}_t, \tilde{e}_t \rangle} \langle \tilde{v}, \tilde{e}_t \rangle \tilde{e}_t
\end{aligned}$$

Least Squares

When represented as a linear system, the task of least squares is to minimize the value of $\|\tilde{y} - A\tilde{u}\|$, which is equivalent to minimizing the sum of the squared residuals. This can be solved by the orthogonal projection of y onto W, where W is the column space of A. This orthogonal projection will be equal to $A\bar{u}$, which simply represents some vector in the column space.

$$\min SSE = \min\|\tilde{y} - A\tilde{u}\| = \tilde{y} - A\bar{u}$$

As $\tilde{y} - A\bar{u}$ is orthogonal to A (it is outside W), it must be the case that:

$$\begin{aligned} A \cdot (\tilde{y} - A\bar{u}) &= 0 \\ &= A^T(\tilde{y} - A\bar{u}) \\ &= A^T\tilde{y} - A^T A\bar{u} \\ A^T\tilde{y} &= A^T A\bar{u} \end{aligned}$$

Note that this final sequence of steps is just a trick to be able to solve for $\bar{u}$. It is possible because we know that, being related to the orthogonal projection of $\tilde{y}$ onto W, $A \cdot (\tilde{y} - A\bar{u})$ is the only dot product in this vector space that will equal zero. Other vectors are not perpendicular to W.

$\hat{y}$: the vector of dependent variable observations
$\bar{u}$: the vector of coefficients of the line of best fit
A: the matrix with a first column of ones, and one column for each vector of independent variables

To construct a correlation matrix:

- Subtract the average of each series from each value in that series
- Place the resulting data down columns in a matrix A_0
- Divide each column of the matrix by its norm to yield A_1
- The correlation matrix will be equal to $A_1^T A_1$

Gram-Schmidt Algorithm

The Gram-Schmidt Algorithm is a method to convert any basis into an orthonormal basis.

$$\text{take } \{\tilde{b}_1, \tilde{b}_2, \tilde{b}_3\} \text{ as the initial basis}$$

$$\begin{aligned} v_1 &= u_1 = \frac{b_1}{\|b_1\|} \\ v_2 &= \frac{u_2}{\|u_2\|} = \frac{b_2 - \langle b_2, v_1\rangle v_1}{\|b_2 - \langle b_2, v_1\rangle v_1\|} \\ v_3 &= \frac{u_3}{\|u_3\|} = \frac{b_3 - \langle b_3, v_1\rangle v_1 - \langle b_3, v_2\rangle v_2}{\|b_3 - \langle b_3, v_1\rangle v_1 - \langle b_3, v_2\rangle v_2\|} \end{aligned}$$

This pattern is extended for every additional vector added to the basis.

Linear Transformations

Basics

A linear transformation is a function that maps between two vector spaces, and in doing so preserves the operations of vector addition and scalar multiplication.

A given linear transformation T can be computed by: $A_T \begin{bmatrix} x \\ y \end{bmatrix}$, where each column of A_T acts on a different basis vector of $\begin{bmatrix} x \\ y \end{bmatrix}$.

Types of Transformations

Reflection in the y-axis: $A^T = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$

Reflection in the line $y = x$: $A^T = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

Rotation by θ degrees anti-clockwise: $A^T = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$

Stretching by a factor c in the x-direction: $A^T = \begin{bmatrix} c & 0 \\ 0 & 1 \end{bmatrix}$

Shear by a factor c along the x-axis: $A^T = \begin{bmatrix} 1 & c \\ 0 & 1 \end{bmatrix}$

Projection onto the y-axis: $A^T = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$

Translation a units to the right and b units up: $A^T = \begin{bmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 1 \end{bmatrix}$

Note that in the case of translation it is necessary to add a dummy coordinate of 1 at the end of all vectors to be transformed.

Image

The image of a linear transformation is the set of all points that can be reached through the linear transformation. It is equal to the span of the column space of the transition matrix A^T.

$$\text{Dim(Im)} = \text{Rank}$$

To determine if a given vector is in the image of a particular transformation, augment the transition matrix with the vector and convert to RREF. If the resulting matrix is inconsistent, that vector is not in the image.

Kernel

The kernel is the set of all vectors that map to the origin under the given linear transformation. That is, it is the set of all vectors that satisfy the equation:

$$T\tilde{u} = \tilde{0}$$

The basis for the kernel is thus the same as the basis for the solution space of the transformation matrix.

$$\text{Dim(Ker)} = \text{Dim(Solution space)}$$

Change of Basis

Written in terms of standard bases: $\begin{bmatrix} x \\ y \end{bmatrix}_S = x(1,0) + y(0,1)$

Written in terms of an alternate basis B: $\begin{bmatrix} x \\ y \end{bmatrix}_B = x\tilde{b}_1 + y\tilde{b}_2$

The matrix that converts from basis B to standard basis S is given by: $P_{S,B} = \begin{bmatrix} \vdots & \vdots & \vdots \\ \tilde{b}_1 & \tilde{b}_2 & \tilde{b}_n \\ \vdots & \vdots & \vdots \end{bmatrix}$

To convert back, simply take the inverse: $P_{S,B} = P_{B,S}^{-1}$

Transformations and Bases

Linear transformations can also act upon vectors written in alternative bases, using this formula:

$$[T]_{C,B} = P_{C,S}[T]_S P_{S,B}$$

$[T]_S$: the transformation matrix in terms of standard bases
$P_{S,B}$: transition matrix from basis B to standard bases
$P_{C,S}$: transition matrix from standard bases to basis C

A transformation matrix $[T]_{B,B}$ gives the output to each of the basis vectors in B as its columns. The entries in each row indicate how many of the corresponding basis vector goes in to form that output for that inputted basis vector. Thus, the following matrix outputs $b_1 + b_2$ when b_1 is inputted, b_2 when b_2 is inputted, and $-2b_1$ when b_3 is inputted.

$$[T]_{B,B} = \begin{bmatrix} 1 & 0 & -2 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Also note that transformations can be given in numerous forms of notation:

$$[T(x,y)]_S = A_T[(x,y)]_S = [T]_{S,S}[(x,y)]_S = A_T \begin{bmatrix} x \\ y \end{bmatrix}$$

Eigenvalues and Eigenvectors

Basics

Eigenvalues and eigenvectors are properties of a given square matrix A. Eigenvectors $\tilde{v}$ are non-zero vectors that, after being multiplied by A, remain parallel to the original vector. For each eigenvector, the corresponding eigenvalue λ is the scalar factor by which the eigenvector is scaled when multiplied by the matrix. This can be represented as follows:

$$A\tilde{v} = \lambda\tilde{v}$$

The scalar λ is said to be the eigenvalue of A corresponding to eigenvector $\tilde{v}$.

Finding Eigenvalues

The basic eigenvector equation can be rewritten as:

$$\begin{aligned} A\tilde{v} &= \lambda\tilde{v} \\ A\tilde{v} &= \lambda I\tilde{v} \\ A\tilde{v} - \lambda I\tilde{v} &= \tilde{0} \\ (A - \lambda I)\tilde{v} &= \tilde{0} \end{aligned}$$

For this to have a non-zero solution, it is necessary that $\text{Rank}(A) < n$. This is the same as saying $\det(A) = 0$. Any non-zero vectors that satisfy this equation will be eigenvectors.

The diagonal entries of a triangular matrix will be equal to that matrix's eigenvalues.

Finding Eigenvectors

Having found the eigenvalues of a given matrix A, finding the corresponding eigenvectors is a simple matter of solving the linear system:

$$(A - \lambda I)\tilde{v} = \tilde{0}$$

$$[A - \lambda I \quad | \quad \tilde{0}]$$

Each system will yield a vector solution $\tilde{v}$, which will be the eigenvector corresponding to the inputted eigenvalue λ.

Diagonalization

Suppose that $\tilde{u}$ and $\tilde{v}$ are eigenvectors of matrix A, with λ_1 and λ_2 being the eigenvalues.

$$A = PDP^{-1}$$

$$A = \begin{bmatrix} v_1 & u_1 \\ v_2 & u_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} u_2 & -u_1 \\ -v_2 & v_1 \end{bmatrix}$$

This can be interpreted as a linear transformation acting on the basis vectors u and v:

$$\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = P^{-1}AP$$

$$[A]_U = P_{U,S}[A]_S P_{S,U}$$

Here $[A]_U$ is the transformation corresponding to A in the basis U, it multiplies by λ_1 in the direction of u and multiplies by λ_2 in the direction of v.

An $n \times n$ matrix can only be diagonalized if it has n linearly independent eigenvectors, which will occur automatically if it has n distinct eigenvalues. Otherwise the independence needs to be checked manually.

Symmetric Matrices

The matrix of eigenvalues Q of a symmetric matrix A has the special property that $Q^{-1} = Q^T$. With this knowledge it is possible to prove that, if a symmetric matrix has positive, non-zero eigenvalues, then it will represent a potential inner product.

Powers of a Matrix

If A is a diagonalizable matrix, then:

$$A^k = PD^kP^{-1}$$

$$A^k = \begin{bmatrix} v_1 & u_1 \\ v_2 & u_2 \end{bmatrix} \begin{bmatrix} \lambda_1^k & 0 \\ 0 & \lambda_2^k \end{bmatrix} \begin{bmatrix} u_2 & -u_1 \\ -v_2 & v_1 \end{bmatrix}$$

Eigenspace: the set of all eigenvectors for a given eigenvalue