Linear Algebra

Paul Kang

December 10, 2020

Contents

1 Vector Space

1.1 Space and Subspace

(Abstract) Vector Space
Any set V can be a space as long as it fulfills the following conditions:
• the space is closed under addition; if v_1, v_2 are in V, then so is $v_1 + v_2$
• the space is closed under scalar multiplication; if c is scalar, then cV is in V
• there exists a zero vector such that $\vec{0} + \vec{v} = \vec{v}$
• for every v, there is a unique additive inverse such that $-\vec{v} + \vec{v} = 0$
• parenthetical algebra is conserved

Examples of common spaces are $\mathbb{R}^{n \times m}$ (space containing n x m matrices containing real numbers), $\mathbb{P}^n$ (space containing polynomials of degree n), $\mathbb{Q}^m$ (space containing rational numbers of degree m).

Subset / Subspace
If V is a vector space, then S is a subspace of V if:
• S is not empty
• $\vec{0}$ exists in S
• S is closed under vector addition
• S is closed under scalar multiplication
Keep an eye out for what is the defining characteristic of the subspace and apply analysis to it and make sure that elements of S are indeed in V.

Span
An element, u, is said to be in the *span* of space V if u can be produced as a linear combination of the elements within V. This stems from the fact that if $u, v \subset V$ then so is $u + v \subset V$. If W is a subspace of space V, and S is a set of vectors

$$c_1\vec{v_1} + c_2\vec{v_2} = 0 \tag{5}$$

Figure 2: The Following Equation must only have a Trivial Solution for the Vectors to be Considered Linearly Independent

within W, then it is said that S spans W or $W = span(S)$

Basis & Dimension

A subset, B, of vector space V is considered a *basis* for V if B spans V and is linearly independent. If B is a subspace and only one of two properties are fulfilled then the basis can still be found:

• if V = span(U) then U contains the basis; find it by removing elements from U

• if U is linearly independent then U is already contained within the basis; find it by adding vectors until V = span(U).

The basis of a vector space is, in essence, the key features in that space such that any element can be made with a linear combination of the basis.

The *dimension* of a vector space V is equal to the sum of the lengths of the vectors in its basis.

$$dim(P) = \infty \tag{1}$$
$$dim(P^n) = n + 1 \tag{2}$$
$$dim(R^{2\times 2}) = 4 \tag{3}$$
$$dim(\{\vec{0}\}) = 0 \tag{4}$$

Figure 1: Sample Dimensions of Spaces

2 Linear Independence and Transformations

2.1 Properties

Orthogonality

A set of vectors are considered *orthogonal* if the dot product of the said vectors equal to 0. This applies to spaces as well. If S is the space of interest, then $S^\perp$ is considered to be the orthogonal complement. For spaces, it may be difficult to check every element with $S^\perp$, but all that really needs to be checked is the basis of S.

$$\|\vec{v}\|^2 = x^2 + y^2 + z^2 \tag{6}$$
$$\|\vec{v} + \vec{u}\|^2 = (\vec{v} + \vec{u})^2 \tag{7}$$
$$= \|\vec{v}\|^2 + 2(\vec{v} \cdot \vec{u}) + \|\vec{u}\|^2 \tag{8}$$
$$\vec{w} = \frac{\vec{v}}{\|\vec{v}\|} \tag{9}$$

Figure 3: Determining the Norm of a Vector and Normalization of the Vector

The bare essential basis known as the *orthonormal* basis set can be found by normalizing the found orthogonal basis set.

Projection

Projection of A onto C is B $\equiv B = proj_C(A)$

Vectors can project their components onto any other defined vector or space with the following equation:

$$proj_{\vec{v}}\vec{u} = (\frac{\vec{u} \cdot \vec{v}}{\|\vec{v}\|^2})\vec{v} \tag{10}$$
$$proj_S\vec{u} = \sum_i proj_{\vec{S_i}}\vec{u} \tag{11}$$

Figure 4: Projection Equation

Always check for orthogonality of the basis of space S with that of the projecting vector to confirm that an *orthogonal* basis set is being used.

The *Gram-Schmidt* algorithm is one that confirms the existence of an orthogonal basis set ('obasis') of space S given basis vectors.

$$basis(S) = span\{\vec{s_1} \dots \vec{s_k}\} \tag{12}$$
$$\vec{v_1} = \vec{s_1} \tag{13}$$
$$\vec{v_2} = \vec{s_2} - proj_{\vec{v_1}}\vec{s_2} \tag{14}$$
$$\vec{v_3} = \vec{s_3} - proj_{\vec{v_1}}\vec{s_3} - proj_{\vec{v_2}}\vec{s_3} \tag{15}$$
$$\vdots \tag{16}$$
$$\vec{v_k} = \vec{s_k} - \sum_{k=1} proj_{\vec{v_{k-1}}}\vec{s_k} \tag{17}$$
$$\therefore obasis(S) = span\{\vec{v_1} \dots \vec{v_k}\} \tag{18}$$
$$\tag{19}$$
$$span\{\vec{v_1} \dots \vec{v_k}\} = span\{\vec{s_1} \dots \vec{s_k}\} \tag{20}$$

Figure 5: Gram-Schmidt Algorithm

2.2 Solving for Linear Systems

<u>'A = LU' Factorization</u>
Easiest Way to Solve for Linear Systems

$$A\vec{x} = \vec{b} \tag{21}$$
$$A = LU \tag{22}$$
$$\begin{cases} L\vec{y} = \vec{b} \\ U\vec{x} = \vec{y} \end{cases} \tag{23}$$

Figure 6: Best Method to Solve Matrix Transformations

$$A = \begin{bmatrix} 2 & 1 & 1 & 4 \\ 7 & 8 & 5 & 5 \\ 1 & 2 & 1 & 9 \\ 8 & 7 & 5 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -7/2 & 1 & 0 & 0 \\ -1/2 & -1/3 & 1 & 0 \\ -4 & -2/3 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 & 4 \\ 0 & 9/2 & 3/2 & -9 \\ 0 & 0 & 0 & 10 \\ 0 & 0 & 0 & 0 \end{bmatrix} = LU \tag{24}$$

Figure 7: Sample Breakdown

The L matrix is lower triangular and is the *record* of Gaussian eliminations applied to the starting matrix A to obtain the upper triangular row reduced echleon matrix U. The columns of L (going left to right) show the row reduction steps used for Gaussian elimination where the 1st column shows the manipulation of the 1st row applied to every other row.
In the case of a rectangular matrix

Row interchange is usually not needed in order to achieve LU factorization. However, in the scenario where it is required, it can be done by applying such operations in the form of elementary matrices. This concept is related to why L above behaves as a "record" for the Gaussian eliminations applied.

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 2 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} = E_2 E_1 A \tag{25}$$

Figure 8: Applying a Row Interchange and Gaussian Elimination with Elementary Matrices

•An important characteristic to note is the order the elementary matrices are applied to the matrix A. The Gaussian elimination was applied first but it appears second and this order is due to matrix algebra mumbo jumbo. Just watch for it. By this example it can be seen that the L matrix is just a product of a series of elementary matrices, making it a "record" to achieve matrix U from A.

'A = QR' Factorization

$$A = QR \tag{26}$$

$$Q^T A = R \tag{27}$$

Figure 9: Basic Process

$$A = Q_1 R_1 \tag{28}$$

$$R_1 Q_1 = A_1 = Q_2 R_2 \tag{29}$$

$$R_2 Q_2 = A_2 = Q_3 R_3 \dots \tag{30}$$

$$R_N Q_N = E \tag{31}$$

Figure 10: Obtaining Eigenvalues

This method can be used to solve for the eigenvalues of the system through diagonalization. Q represents a matrix consisting of the orthonormal basis vectors of A as the columns of Q. The R matrix will be a square matrix that is upper triangular with positive diagonal entries. The Q matrix can be determined by taking the columns of A and applying the Gram-Schmidt algorithm to generate the orthonormal columns of Q. The matrix R will be a square upper triangular matrix with positive diagonal entries. Repeating the process by switching the positions of Q and R until no new matrices are generated will result in matrix E, which will contain the eigenvalues of matrix A in the diagonals.

Least Squares Regressions
The regression minimizes the distance between the predicted value and the projection of the real value onto the approximation. This can be achieved by solving for the $\vec{x}$ with:

3 Linear Transformations

Linearity
Properties:

$$A^T A\vec{x} = A^T \vec{y} \tag{32}$$
$$R\vec{x} = Q^T \vec{y} \tag{33}$$

Figure 11: The Derived Formula for Regression; the Second is Preferred if QR Factorization Provided

•$T(\vec{v_1} + \vec{v_2}) = T(\vec{v_1}) + T(\vec{v_2})$
•$T(c\vec{v}) = cT(\vec{v})$

3.1 Eigenvectors and Eigenvalues

$$A\vec{x} = \lambda\vec{x} \tag{34}$$

Figure 12: Eigenvalue, Eigenvectors and Eigenspace

Definition
A linear transformation involving a square matrix (is it specific to square matrices?) applied to an eigenvector such that this would be the same as a scalar eigenvalue applied to the same eigenvector. Eigenvectors associated with unique eigenvalues are independant and orthogonal. If eigenvectors share the same eigenvalue, then the share the same space. What is interesting is that the "capability" of the transformation can be found within the transformation matrix itself.

Finding Eigenvectors and Eigenvalues

$$det(A - \lambda I) = 0 \tag{35}$$
$$(A - \lambda_x I)u = 0 \tag{36}$$
$$u = eigenvector_x \tag{37}$$

Figure 13: Solving for Eigenvocabulary

•Step 1: Determine the eigenvalues associated with the matrix by finding the characteristic polynomial from the 1st equation. If there are two eigenvalues that are the same here, then there must be two associated eigenvectors.
•Step 2: Find the associated eigenvectors with by solving for the second equation and finding what vector, u, makes the system zero out. If the eigenvalue is degenerate, then free variables will appear; i.e. if there are two degenerate eigenvalues, then there will be two free variable vectors that are the associated eigenvectors. The set of eigenvectors within the associated eigenvalue form the basis for the associated *eigenspace*.

$A = PDP^{-1}$
•A: the *square* transformation matrix
•D: λI; a matrix where the eigenvalues line the diagonals
•P: a matrix where every column is an orthonormalized eigenvector corresponding with an eigenvalue from D, i.e. eigenvalue of row 1 column 1 corresponds with the orthonormal column 1 of P (all columns are independant). If the dimension of the matrix consisting of the eigenvectors is less than that of the starting matrix, A, then the matrix is *not* diagonizable.
•If an eigenvalue is degenerate, then corresponding eigenvectors must be orthonoramlized

3.2 Properties

$$dim(\text{input space}) = dim(ker(T)) + dim(range(T)) \tag{38}$$

Figure 14: Rank-Nullility Theorem

Linear transformations can be observed indirectly through the transformation itself in order to better understand what is happening from the input space (*domain*) into the output space (*codomain*).

Kernal
•is a subspace of the domain
•contains all the vectors from the domain that point to the zero vector in the codomain

$$T(\vec{x}) = 0 \tag{39}$$

Figure 15: Solving for the Kernal of T

Range
•is a subspace of the codomain
•the range of T contains all possible images or outputs
•Rank-Nullility theorem can be used to find a "skeleton" for the basis of $range(T)$

3.3 Types of Transformations

One-to-One
•1 input has 1 or more outputs
•the detail: all output can be traced back to *a unique* 1 input
•if $ker(T)$ is only contains the zero vector, then the *set is of dimension 0* and makes the transformation 1-to-1

Onto
•multiple inputs have 1 output
•the detail: every output can be traced back to *at least* 1 input
•if $range(T) =$ output, then the transformation is onto

Isomorphism
•if a transformation is both 1-to-1 and onto (i.e. has 1 input →1 output), then the input and output space are the same and the transformation is considered isomorphic
•if there is a T that is isomorphic that takes an input and forms an output, then the inverse, T^{-1}, can be applied to revert the output back into the input. This inversibility is *unique* to isomorphic transformations.

3.4 Coordinate Vectors and the Matrix of Linear Transformation

Coordinate Vectors
A way to express large abstract vectors in a simpler form

$$B = \{\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}1\\-1\end{bmatrix}\} \tag{40}$$

$$\vec{v} = \begin{bmatrix}2\\3\end{bmatrix} = x\begin{bmatrix}1\\1\end{bmatrix} + y\begin{bmatrix}1\\-1\end{bmatrix} \tag{41}$$

$$[\vec{v}]_B = \begin{bmatrix}5/2\\-1/2\end{bmatrix} \tag{42}$$

Figure 16: Sample of Determining Coordinate Vector relative to B

•The subscript of the coordinate vector will specify what basis the vector was relative to and will be different based on the basis used
•the coordinate vector represents the values needed to create a vector relative to a basis
•coordinate vectors are another way of expressing transformations; determine the coordinate vector in terms of the output space and apply the vector to the basis of the output space to optain the image.

Creating the Matrix of Linear Transformation

$$A[\vec{v}]_U = [T(\vec{v}]_B \tag{43}$$

Figure 17: Basic Premise; U represents the basis of the input space and B represents the basis of the output space

$$T : R^3 \to P^2,\ T(\begin{bmatrix} a \\ b \\ c \end{bmatrix}) = (c-b) - 2bx + (b-a)x^2 \tag{44}$$

$$\text{basis of input space} = U = \{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\},\ \text{basis of output space} = V = \{1, x, x^2\} \tag{45}$$

$$T(\vec{u_1}) = -1 - 2x + x^2 \to [T(\vec{u_1})]_B = \begin{bmatrix} -1 \\ -2 \\ 1 \end{bmatrix} \tag{46}$$

$$A = \begin{bmatrix} 0 & -1 & 1 \\ 0 & -2 & 0 \\ -1 & 1 & 0 \end{bmatrix} \tag{47}$$

Figure 18: Sample Procedure

•Step 1: Apply transformation to basis of *input* space
•Step 2: Find the coordinate vectors of the images relative to the *output space*
•Step 3: Following the order of the basis columns, combine the coordinate vectors into one matrix to create the matrix of linear transformation

3.5 Change of Basis

Basis Transformation
A special type of linear transformation is the conversion of a vector relative to one basis into a new vector relative to another basis *within the same space*. The process of finding the matrix of this unique transformation is similar to process described in the previous section.

•Step 1: Apply transformation to the *output basis*
•Step 2: Find coordinate vectors of the images relative to the *output basis*
•Step 3: Following the order of the basis columns, combine the coordinate vectors into one matrix to create a matrix that converts output space into input space
•Step 4: determine inverse to regain direction from input to output

Similarity

$$B = S^{-1}AS \tag{48}$$

$$A = SBS^{-1} \tag{49}$$

Figure 19: Similarity Relationship

For the following: $A \equiv$ transformation matrix relative to one basis, $B \equiv$ transformation matrix relative to another basis and $S \equiv$ similarity matrix, if the relationship defined above holds true, then matrix A is deemed *similar* to matrix B. This form is similar to the diagonalization of a linear square matrix into eigenvalues and eigenvectors. This implies that the transformation matrix is converted into a new transformation matrix consistent of eigenvalues relative to the basis defined by the eigenvectors.

4 Inner Product

Definition & Properties

$$< \vec{u}, \vec{v} >_A = \vec{u}^T A^T A \vec{v} = (A\vec{u}) \cdot (A\vec{v}) \tag{50}$$
$$< A, B > = tr(A^T B) \tag{51}$$

Figure 20: Inner Product Notation with Respect to Matrix A

$$< \vec{u} + \vec{v}, \vec{w} > =< \vec{u}, \vec{w} > + < \vec{v}, \vec{w} > \tag{52}$$
$$< c\vec{u}, \vec{v} > = c < \vec{u}, \vec{v} > \tag{53}$$
$$< \vec{u}, \vec{u} > = ||A\vec{u}||^2 \tag{54}$$
$$\tag{55}$$

Figure 21: Inner Product Properties

$$< f(x), g(x) >= \int f(x)g(x)dx \tag{56}$$

Figure 22: Equation Relevance

$$proj_B A = \frac{< A, B >}{< B, B >} B \tag{57}$$

Figure 23: Space Projection

•The inner product is just an extension of the dot product and follows the same properties.
•if matrices are involved, then A must have linearly independant columns
Fourier Approximation

$$< f, g >= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)g(x)dx \tag{58}$$

Figure 24: Fourier Equation over Inner Product Space $V = [-\pi, \pi]$

•used trigonometric functions to approximate periodic functions; useful for projecting an unknown function onto a periodic function

4.1 Quadratic Form

Notation

$$5x_1^2 + 4x_1x_2 + 8x_2^2 = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 5 & 2 \\ 2 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{59}$$

Figure 25: Sample Notation

•Expressing quadratic equations as matrices (because every constant by a variable can be called from a matrix

$$\vec{x}^T A \vec{x} = \vec{y}^T D \vec{y} \tag{60}$$
$$\vec{y} = P^T \vec{x} \tag{61}$$

Figure 26: Principal Axes Theorem where A is the matrix of transformation, D is the diagonal matrix of eigenvalues and P is the eigenvector matrix

•any quadratic form can be rewritten with a matrix of eigenvalues (derived from A) and change in variables to form $\vec{y}$

Types of Quadradic Forms

•Positive *Definite* if $Q(\vec{x}) > 0$ for all $\vec{x} \neq \vec{0}$, also if all eigenvalues are *positive*; *Semidefinite* if for all $\vec{x}$
•Negative *Definte* if $Q(\vec{x}) < 0$ for all $\vec{x} \neq \vec{0}$, also if all eigenvalues are *negative*; *Semidefinite* if for all $\vec{x}$
•*Indefinite* if a bit of everything