Linear Algebra Notes

Chris Beard

28 July 2011

Contents

1 Linear Algebra Notes

1.1 Basic Linear Algebra concepts

13. Singular matrix, no inverse for matrix M; $|M| = 0 \Leftrightarrow M^{-1}$ undefined 13

14. Line: $L = \{\vec{x} + t\vec{v} | t \in \mathbb{R}\}$

18. Set of n linear independent vectors $\mathcal{V} = \{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\}$:
    - $c_1\vec{v}_1 + c_2\vec{v}_2 + ... + c_n\vec{v}_n = 0 \Rightarrow c_1 = c_2 = ... = c_n = 0$
    - otherwise, $\mathcal{V}$ is linearly dependent

19. Subspace: vector space
    - Span of n vectors is valid subspace of $\mathbb{R}^n$
    - Subspace must be closed under addition
    - and scalar multiplication
      - must include $\{\vec{0}\}$

20. Basis: minimum set of linearly independent vectors that spans a subspace

21. Dot product: $\vec{a} \cdot \vec{v} = a_1 v_1 + a_2 v_2 + ... + a_n v_n$
    - commutative
    - distributive
    - associative
23. Cauchy-Schwartz inequality
    - $|\vec{x} \cdot \vec{y}| \leq \|\vec{x}\| \|\vec{y}\|$
    - special case: $\vec{x} = c\vec{y} \Leftrightarrow |\vec{x} \cdot \vec{y}| = \|\vec{x}\| \|\vec{y}\|$
24. Vector triangle inequality
    - $|\vec{x} + \vec{y}| \leq \|\vec{x}\| + \|\vec{y}\|$
    - special case: $|\vec{x} + \vec{y}| = \|\vec{x}\| + \|\vec{y}\| \Leftrightarrow \vec{x} = c\vec{y}$
    - simple, but helpful for scaling to general coordinates

1.2 Geometry

1. Angle θ between vectors
    - $\vec{a} \cdot \vec{b} = \|\vec{a}\| \|\vec{b}\| \cos\theta$
    - $\vec{a} \perp \vec{b} \Rightarrow \vec{a} \cdot \vec{b} = 0$
2. Find a plane
    - $\vec{x}_0$ is a point on the plane, $\vec{n}$ is a vector normal to the plane, $\vec{x}$ is a vector that ends on the plane
    - get plane in traditional form: $Ax + By + Cz = D$
        - via: $n_1(x - x_0) + n_2(y - y_0) + n_3(z - z_0) = 0$

29. Area of a parallelogram formed by vectors $\vec{a}, \vec{b}$
    - $A = \|\vec{a} \times \vec{b}\|$

1.3 Nullspace $\mathcal{N}$

34. Subspace: Nullspace $N = \{\vec{x} \in \mathbb{R}^n | A\vec{x} = \vec{0}\}$
    - $N(A) = N(rref(A))$
    - $N(A) = 0 \Leftrightarrow$ column vectors are linearly independent
36. $\mathcal{N}(A) = \{0\} \Leftrightarrow$ column vectors are linearly independent
40. Any basis of a given span will have the same number of elements
42. $\mathbf{dim}(C(A)) = \mathbf{rank}(A)$ (for column space of $A, C(A)$)

1.4 Transformations

47. Linear transformations of x can be represented as Ax for some matrix A

    - linearity:
        - $A(x+y) = Ax + Ay$
        - $A(\alpha x) = \alpha Ax$

48. Matrix product with any vector is a linear transformation

    - any linear transformation can be represented as a matrix vector product

51. $\mathcal{V}$ a subspace in $\mathbb{R}^n$

    - $T : \mathbb{R}^n \to \mathbb{R}^m$
        - $T(\mathcal{V})$ is image of $\mathcal{V}$ under T
        - $T(\mathbb{R}^n)$ is image of actual 'T'

52. $T : X \to Y$, S is a subset of domain Y

    - everything in $x \subseteq X$ maps to some place in Y
    - everything in $S \subseteq Y$ gets mapped to
    - $T^{-1}(S)$ is pre-image of S under T
    - $T^{-1}(S) = \{\vec{x} \in \mathbb{R}^n | T(\vec{x}) \in S\}$
    - **kernel**$(T) = \{\vec{x} \in \mathbb{R}^n | T(\vec{x}) = \{0\}\}$

57,58. Rotation transformation

- rotation is a linear $\Leftrightarrow$ representable as matrix product.
- Basic method for $\mathbb{R}^3$: represent rotation around x axis with matrix A, then about y axis with a matrix B, then about z axis with matrix C. (do this on unit vectors to find A, B, C). The resulting transform on a vector $\vec{x}$ will be $C(B(A\vec{x}))$
- In $\mathbb{R}^2$:

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

- In $\mathbb{R}^3$, around x-axis:

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{bmatrix}$$

60. Projections of $\vec{x}$ onto space $L = \{c\vec{v} | c \in \mathbb{R}\}$

    - $\mathbf{Proj}_L(\vec{x}) = \left(\frac{\vec{x}\cdot\vec{v}}{\vec{v}\cdot\vec{v}}\right)\vec{v} = \frac{\vec{x}\cdot\vec{v}}{\|\vec{v}\|^2}\vec{v}$
    - easier with unit vector

61. Projections are easier if you have a normalized vector $\vec{u} = \frac{\vec{v}}{\|\vec{v}\|}$:

    - $\boxed{\mathbf{Proj}_L(\vec{x}) = (\vec{x}\cdot\vec{u})\,\vec{u}} = A\vec{x}$, where $\vec{u}$ is a unit vector on L
    - if $\hat{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$, get $A_{n\times n}$ by applying the transformation $\mathbf{Proj}_L(\vec{x})$ to each column in I_n

63. Given $T(\vec{x}) = B_{l\times m}\vec{x}, S(\vec{x}) = A_{m\times n}\vec{x}$
    - Then T composed with S is given as $T \circ S(\vec{x}) = T(S(\vec{x})) = B(A\vec{x})$
    - Remember interpretation of matrix vector products:

$$A\vec{x} = \begin{bmatrix}\vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n\end{bmatrix}\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix} = x_1\vec{a}_1 + x_2\vec{a}_2 + ... + x_n\vec{a}_n$$

    - In other words, the product of a matrix A and a vector $\vec{x}$ is a $\boxed{\text{linear combination of the column matrices}}$ of A, $\vec{a}_i$, scaled by the elements of $\vec{x}$, x_i

1.5 Invertibility

67. Function f is invertible $\Leftrightarrow \exists$ unique inverse function of f
68. Invertibility $\Leftrightarrow$ unique solution
69. Surjective (onto) and injective (one-to-one)
    - Surjective: $\forall y \in Y \exists$ at least one $x \in X : f(x) = y$
      - Everything in the co-domain gets mapped to, everything reachable
    - Injective: for any $y \in Y, \exists$ at most 1 $x : f(x) = y$
      - One-to-one correspondence
70. Invertibility $\Leftrightarrow$ injective (one-to-one)
71. Transformation $T : \mathbb{R}^n \to \mathbb{R}^m, T(\vec{x}) = A\vec{x}$
    - T onto (surjective) $\Leftrightarrow \textbf{span}(C(A)) = \mathbb{R}^m$
72. Viewing a plotted solution set of $Ax = b$ in $\mathbb{R}^2$
    - solution set is a shifted version of the nullspace, assuming there is a solution
    - Solution set is particular + homogenous solution: $\{\vec{x}_p\} + \mathcal{N}(A)$
73. Requirements for A representing an injective (1-1) transformation:
    - $\mathcal{N}(A) = \{0\}$, trivial nullspace, which implies the following:
      - Column vectors of A are LI
      - $C(A) = \textbf{span}(a_i, a_2, ..., a_n)$
      - The column vectors form a basis for $\mathbb{R}^n$
      - rank$(A)= n$
75. Inverses are representable as linear operations
    - $cT(x) = T(cx)$, where T is a linear transformation that gives the inverse
76. Get the inverse of $A \in \mathbb{R}^{n\times n}$ via elementary row operations on matrix augmented with I_n:
    - $[A|I] \to [I|A^{-1}]$
78. Formula for inverse of $A = \begin{bmatrix}a & b\\c & d\end{bmatrix}$
    - $A^{-1} = \frac{1}{ad-bc}\begin{bmatrix}d & -b\\-c & a\end{bmatrix}$

1.6 Determinants

79. Method for 3×3 determinant

80. For $n \times n$ determinant

    - $|A| = \sum_{j=i}^{j=n} (-1)^{i+j} a_{ij} |A_{ij}|$, where A_{ij} is the submatrix formed by eliminating the ith row and jth column

82. Rule of Sarrus of Determinants

83. Effect of scalar multiplication on the determinant

    - for $n \times n$ matrices, $|kA| = k^n |A|$
    - For a single row of A multiplied by k, det is $k|A|$

85. When you add a row from matrix X to a row from Y to get a matrix Z:

    - $|Z| = |X| + |Y|$
    - S_{ij} is A with 2 rows swapped: $|S_{ij}| = -|A|$

86. If A has duplicate rows, $|A| = 0$

87. Adding scaled rows of a matrix to other rows within the matrix does **not** change the determinant

88. Determinant of upper triangular matrix is the product of the diagonal

90. Area of the parallelogram formed by 2 column vectors in A is equal to $|A|$

91. A is the area of some set plotted in $\mathbb{R}^2$; the area of the set after transformed by matrix B is $|\mathbf{det}(B)A|$

    - From Wikipedia, "A 2×2 matrix with determinant -2, when applied to a region of the plane with finite area, will transform that region into one with twice the area, while reversing its orientation."

93. $|A^T| = |A|, A \in \mathbb{R}^{n \times n}$

1.7 Transpose, left nullspace $\mathcal{N}(A^T)$ and rowspace $C(A^T)$

94. Transpose of sum: $C = A + B \Rightarrow C^T = (A + B)^T = A^T + B^T$

    - $(A^T)^{-1} = (A^{-1})^T$

95. Dot product and transpose (for vectors):

    - $v \cdot w = v^T w$
    - $(Ax) \cdot y = x \cdot (A^T y)$

96. Rowspace of A: $C(A^T)$ (column space of A's transpose)

    - Left nullspace: $\{x \in \mathbb{R}^n : (x)^T A = 0^T\}$

97. Any member of a rowspace of A is orthogonal to any member of the nullspace of A

    - or, $C(A) \perp \mathbf{LeftNullspace}(A)$
    - Where left nullspace is equal to $\mathcal{N}(A^T)$

98. $A = B^T \Rightarrow N(B^T) = C(B)^{\perp}$; left nullspace is the orthogonal complement to the column space

99. Rank (A) = Rank (A^T)

100. Rank (A)+ Nullity $(A) = \mathbf{dim}(V) + \mathbf{dim}(V^\perp) = n$
    - Where A is an $n \times n$ matrix and V is a vector space in $\mathbb{R}^n$
101. $V \cup V^\perp = \{\vec{0}\}$ for a vector space $V \subseteq \mathbb{R}^n$
    - $\mathbf{dim}(V) = k, \mathbf{dim}(V^\perp) = n - k$
    - $\{\vec{v}_1, ..., \vec{v}_k\}$ form basis for $V, \{\vec{w}_1, ..., \vec{w}_{n-k}\}$ form basis for $V^\perp$
    - For $\vec{x} \in \mathbb{R}^n$, $\boxed{\vec{x} = c_1\vec{v}_1 + ... + c_k\vec{v}_k + c_{k+1}\vec{w}_1 + ... + c_n\vec{w}_{n-k}}$
        - any vector in $\mathbb{R}^n$ is a linear combination of the basis vectors of $V \subseteq \mathbb{R}^n$ and $V^\perp$
102. $\vec{x} \in (V^\perp)^\perp \Rightarrow \vec{x} \in V$
103. $\mathcal{N}(A^T)^\perp = C(A); C(A^T)^\perp = \mathcal{N}(A)$
    - Column space is equal to the orthogonal complement of the left nullspace
    - Nullspace is equal to the orthogonal complement of the rowspace
104. #101, 103 $\Rightarrow$ for $x \in \mathbb{R}^n, r_0 \in C(A^T), n_0 \in \mathcal{N}(A) \Rightarrow \boxed{x = r_0 + n_0}$
    - $b \in C(A) \Rightarrow \exists$ unique $r_0 \in C(A^T) : r_0$ is a solution to $Ax = b$, and s.t. r_0 has the minimum length of any solution
    - $\boxed{\text{rowspace} \perp \text{nullspace}}$
105. Geometric interpretation, see notes

1.8 Projections and Least Squares Approximation

106. For $A \in \mathbb{R}^{n\times k}$, $\boxed{\text{any } A^TA \text{ is an invertible square matrix}}$
    - $v \in \mathcal{N}(A^TA) \Rightarrow v \in \mathcal{N}(A), v = \vec{0}$
107. Remember, $\mathbf{Proj}_L(\vec{x}) = \left(\dfrac{\vec{x}\cdot\vec{v}}{\vec{v}\cdot\vec{v}}\right)\vec{v} = \dfrac{\vec{x}\cdot\vec{v}}{\|\vec{v}\|^2}\vec{v}$
108. Geometric intepretation of projections, see notes
109. Projection onto a subspace is a linear transformation:
    - $\boxed{\mathbf{Proj}_V(x) = Ay = A(A^TA)^{-1}A^Tx}$, where A has basis vectors of subspace V as its column vectors
112. $\mathbf{Proj}_V(x)$ is the closest vector to x that lies on subspace V
113. Least squares approximation: see Figure 1
    - remember: $\boxed{\text{Any solution } x \text{ for } Ax = b \text{ must lie on column space of } A, C(A)}$
    - if $\nexists$ solution to $Ax = b$, (i.e., b isn't in $C(A)$) we can still find the closest approximation $\hat{x}$ (or x^* in the figure), which is closest to the column space of A
    - this will naturally be $\mathbf{Proj}_{C(A)}b$
    - remember, this is given as $\hat{x} = \mathbf{Proj}_{C(A)}b = A(A^TA)^{-1}A^Tx$
    - when you look at relation of $C(A)$ to $\mathcal{N}(A)$ in the context of p, you get the 'simplification' that $\boxed{A^TA\hat{x} = A^Tb}$
114. LSA example for intersection of 3 non-intersecting lines: given A, b, want to find $\hat{x}$.
    - Calculate A^TA, A^Tb, and put these matrices' augmented matrix into rref to get values of $\hat{x}$
115. LSA example of fitting lines to data points (finding best fit for parameters m, b in $y = mx + b$ notation)

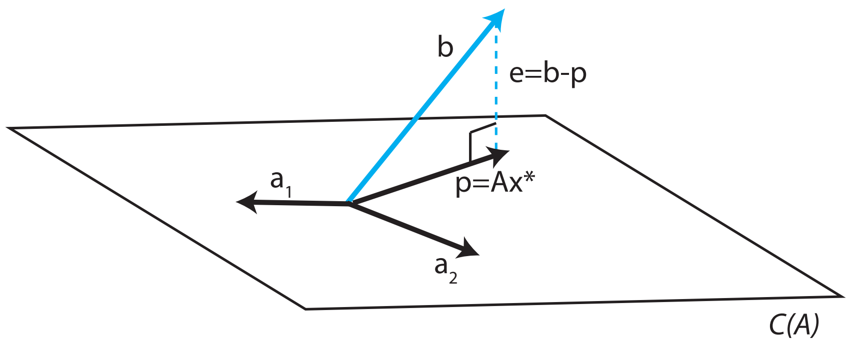

Figure 1: Least square approximation (p) for b that minimizes error e. Plane is the column space of A spanned by a_1 and a_2

1.9 Coordinates in different bases

116. $V \subset \mathbb{R}^n, B = \{v_1, ..., v_k\}$ is basis for V
    - for all $a \in V, a = c_1 v_1 + ... + c_k v_k$
    - $c_1, c_2, \cdots, c_k$ known as 'coordinates with respect to B'
    - $[a]_B = \begin{bmatrix} c_1 & c_2 & \cdots & c_k \end{bmatrix}^T = c$, where scalar elements of c are the weights of bases of V that you need to get a in terms of the basis vectors
    - Example: $[a]_B = \begin{bmatrix} 3 & 2 \end{bmatrix}^T \Rightarrow a = 3v_1 + 2v_2$
    - Standard coordinates for $\mathbb{R}^2$ use $\mathbf{e}_1 = \begin{bmatrix} 1 & 0 \end{bmatrix}^T$, $\mathbf{e}_2 = \begin{bmatrix} 0 & 1 \end{bmatrix}^T$
117. $V \subset \mathbb{R}^n, B = \{v_1, ..., v_k\}$ is basis for V:
    - $C = \begin{bmatrix} v_1 & v_2 & \cdots & v_k \end{bmatrix}^T \in \mathbb{R}^{n \times k}$ is change of basis matrix for subspace V
    - $\boxed{a = C[a]_B}$: the basis vectors (C) multiplied by their weights ($[a]_B$) give the vector a in standard coordinates
118. If C (matrix with bases of V in B as column vectors) is invertible
    - C is square $\Rightarrow k$ (number of basis vectors) $= n$ (number of rows/dimension of basis vectors) $\Rightarrow \exists n$ basis vectors
    - C has Linearly Independent columns
    - B (set of basis vectors $\in \mathbb{R}^n$ for subspace V) is basis for $\mathbb{R}^n$
    - $\boxed{C \text{ invertible} \Leftrightarrow \textbf{span}(B) = \mathbb{R}^n}$
119. Transformation matrix w.r.t. different basis; $T : \mathbb{R}^n \to \mathbb{R}^n$ ($T(x) = Ax$, A transforms wrt standard basis). See Fig 2
    - $\boxed{D = C^{-1}AC}$, where D gets you from vector a in terms of B to transformation of a in terms of B
        - i.e., $D[\vec{x}]_B = [T(\vec{x})]_B$
        - $\boxed{A = CDC^{-1}}$

122. Example: Transformation that reflects across a line $L = \{c\hat{u}_1 : c \in \mathbb{R}\}$ (in $\mathbb{R}^2$)

$$\begin{array}{ccc} \vec{x} & \xrightarrow[CDC^{-1}]{A} & T(\vec{x}) \\ C^{-1}\big\downarrow\big\uparrow C & & C^{-1}\big\downarrow\big\uparrow C \\ [\vec{x}]_B & \xrightarrow[C^{-1}AC]{D} & [T(\vec{x})]_B \end{array}$$

Figure 2: Relationship of coordinate transformation matrices. $\vec{x}$ is a vector in standard coordinates, and $T(\vec{x}) = A\vec{x}$ is the transformation, in standard coordinates. These terms enclosed with brackets indicates the vector with respect to the basis B, whose elements form the columns of C. C gets you from B coordinates to standard coordinates, and the inverse reverses this.

- Use vector $\hat{u}_1 = [1\ 2]^T = [[1\ 0]^T]_B$ and $\hat{u}_2 = [2\ -1]^T = [[0\ 1]^T]_B$ (s.t. $\hat{u}_1 \perp \hat{u}_2$) as new basis vectors, which will be the columns of C
- A will have the transformations of the standard unit vectors $\mathbf{e}_1, \mathbf{e}_2$ as its columns (which are hard to find out)
- D will have the transformations of $\hat{u}_1, \hat{u}_2$ in modified coordinates as its columns
  - this is easy: $\hat{u}_1$ will remain the same, $\hat{u}_2$ will be negated, i.e.,
    * $[T(\hat{u}_1)]_B = [\hat{u}_1]_B = [[1\ 0]^T]_B$
    * $[T(\hat{u}_2)]_B = [-\hat{u}_2]_B = [[0\ -1]^T]_B$
    * These are column vectors of D
- $A = CDC^{-1}$

1.10 Orthonormal bases

123. Orthonormal bases: basis vectors orthogonal and $\perp$ to each other

124. $B = \{v_1, ..., v_k\}$ is ON basis for $V, \vec{x} \in V \Rightarrow \vec{x} = c_1v_1 + \cdots + c_kv_k$

   - if B is orthornormal basis, $v_i \cdot \vec{x} = c_i(v_k \cdot v_i) = c_i$, because of sifting property: $v_i \cdot v_j = \{i \neq j : 0, i = j : 1\}$
   - Remember weights for change of basis matrix

$$[\vec{x}]_B = \begin{bmatrix} c_1 & c_2 & \cdots & c_k \end{bmatrix}^T = \begin{bmatrix} v_1 \cdot x_1 & v_2 \cdot x_2 & \cdots & v_k \cdot x_k \end{bmatrix}$$

   - so $[x]_b$ is easy to find

125. $B = \{v_1, v_2, ..., v_k\}$ is ON basis for subspace $V \subset \mathbb{R}^n$

   - for $x \in \mathbb{R}^n, x = \vec{v} + \vec{w}$, where $\vec{v} = \mathbf{Proj}_V x \in V, \vec{w} = \mathbf{Proj}_{V^\perp} x \in V^\perp$
   - remember, #109: $\boxed{\mathbf{Proj}_V(x) = Ay = A(A^TA)^{-1}A^Tx}$, where A has basis vectors of subspace V as its column vectors
   - if $\boxed{B \text{ is an ON basis}}$, $A^TA = I_k$, because of ON sifting property (i.e., $A^T = A^{-1}$)
   - $\boxed{\mathbf{Proj}_V(x) = AA^Tx}$

128. Orthogonal matrices preserve angles and lengths (i.e., when ON matrix C is used as a transformation matrix)

   (a) Length
      - remember $y \cdot y = y^Ty$
      - with $\|x\|^2 = \|Cx\|^2$, the C disappears because of above property

   (b) Angles
      - look at $\cos\theta$ of 2 vectors from its relation to the dot product

- compare this with θ between vectors Cv, Cw, keeping in mind preservation of length property above

129. Gram-Schmidt process for basis $B = \{v_1, v_2, ..., v_k\}$ for subspace V

- basis not originally orthonormal, but first normalize v_1 (or any vector in the set): $u_1 = v_1/\|v_1\|$
  - $\{u_1\}$ is now an ON basis for $V_1 = \mathbf{span}(v_1) \subset V$
- $V_2 = \mathbf{span}(v_1, v_2) = \mathbf{span}(u_1, v_2) = \mathbf{span}(u_1, y_2)$, where $y_2 = v_2 - \mathbf{Proj}_{V_1}(v_2)$, i.e., y_2 is the element of v_2 that is orthogonal to V_1
  - $y_2 = v_2 - (v_2 \cdot u_1)u_1$; u_2 is normalized version of y_2
  - $\{u_1, u_2\}$ is now ON basis for V_2
- Repeat this for rest of vectors in B
  - But, in general, $y_i = v_i - \mathbf{Proj}_{V_1}(v_i) - \mathbf{Proj}_{V_2}(v_i) - ... - \mathbf{Proj}_{V_{i-1}}(v_i) = v_i - (v_i \cdot u_1)u_1 - (v_i \cdot u_2)u_2 - ... - (v_i \cdot u_{i-1})u_{i-1}$

1.11 Eigenvalues and vectors

132. Eigenvalues; $T : \mathbb{R}^n \to \mathbb{R}^n$

- Remember, from #122, a new basis was chosen for transformation T, s.t. T did not scale basis vectors
  - $T(v_1) = (1)v_1$
  - $T(v_2) = (-1)v_2$
- For any transformation $\boxed{T : T(v) = Av = \lambda v}$
  - v is an eigenvector, and $\lambda \in \mathbb{R}$ is the eigenvalue associated with it

133. Finding λ, s.t. $Av = \lambda v$:

$$\begin{aligned} 0 &= \lambda v - Av && (1) \\ &= \lambda I_k v - Av && (2) \\ &= (\lambda I_k - A)v && (3) \\ \Rightarrow v &\in \mathcal{N}(\lambda I_k - A) = \{x \in \mathbb{R}^n | Bx = 0 \text{ where } (\lambda I_k - A) = B\} && (4) \\ \text{Note: Columns of } D \text{ Linearly independent} &\Leftrightarrow \mathcal{N}(D) = \{0\} && (5) \end{aligned}$$

- So, columns of D must be Linearly Dependent for there to be a non-trivial eigenvector.
  - If columns LD, D not invertible (since invertibility $\Leftrightarrow$ one-to-one/injective)
  - $\boxed{\Rightarrow \mathbf{Det}(\lambda I_k - A) = 0}$ should get you your eigenvalue

134. Characteristic equation (in terms of λ) is found by evaluating the determinant: $\mathbf{Det}\ (\lambda I_k - A) = 0$

135. Find eigenvector: Solve for v in $(\lambda I_n - A)v = 0$

- Won't need to do rref on augmented matrix, since the RHS is the zero vector
- rref $(\lambda I_n - A)$ will give elements of v
- **Eigenspace** will be the span of the eigenvectors

138. Eigenbasis (basis composed of eigenvectors) make good coordinate system

- $A \in \mathbb{R}^{n \times n}$ has n LI eigenvectors, $B = \{v_1, v_2, ..., v_n\}$
  - n vectors that are LI $\Rightarrow B$ is basis for $\mathbb{R}^n$
  - $Av_1 = \lambda_1 v_1 + 0v_2 + ... + 0v_n$; $Av_2 = 0v_1 + \lambda_2 v_2 + ... + 0v_n$ (and you get the idea)
  - Remember chart from #122: vectors in B will form columns of C

– Find column j of D by doing $[T(v_j)]_B = [0\ 0 \cdots \lambda_j \cdots 0]^T = [d_1\ d_2 \cdots d_n][0\ 0 \cdots 1 \cdots 0]^T$

$$\text{Thus, } D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \lambda_n \end{bmatrix}$$

Easy to apply transformation on a whole range of input vectors, and can be worth the overhead of getting the basis and D, if you want to scale up.

2 Appendix

2.1 Why reduced row echelon works

We can use reduced row echelon form to solve a system of equations $Ax = b$. We want to know any of the polynomials of degree $n \leq 2$ that go through the points $\{(1, -1), (2, 3), (3, 3)\}$. We will look for coefficients a_i associated with terms x^i. Matching the x^i values with their respective coefficients and y values gives us the system

$$\begin{aligned} a_0 + a_1 + a_2 &= -1 \quad (6) \\ a_0 + 2a_1 + 4a_2 &= 3 \quad (7) \\ a_0 + 3a_3 + 9a_2 &= 3 \quad (8) \end{aligned}$$

which can be represented in $Ax = b$ form by

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \\ 3 \end{bmatrix}.$$

We can see from Eqns. (6)-(8) that our method should be to perform elementary row operations on the different equations, in order to ultimately have only a single unique a_i term on the LHS of each equation. We want the a_i terms to remain in the same columns, but the coefficients to mostly cancel each other. Thus we will augment matrices A and b, and get them in reduced row echelon form:

$$\textbf{rref}\left(\begin{bmatrix} 1 & 1 & 1 & -1 \\ 1 & 2 & 4 & 3 \\ 1 & 3 & 9 & 3 \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 & 0 & -9 \\ 0 & 1 & 0 & 10 \\ 0 & 0 & 1 & -2 \end{bmatrix}$$

Un-augmenting this matrix, we see that it corresponds to the equations

$$\begin{aligned} a_0 &= -9 \quad (9) \\ a_1 &= 10 \quad (10) \\ a_2 &= -2 \quad (11) \end{aligned}$$

so that the polynomial can be given as $y(x) = -9 + 10x - 2x^2$.

Note, if the last row of the matrix 2.1 had been [0 0 0 1], meaning that $0a_0 + 0a_1 + 0a_2 = 1$, (an obvious contradiction), it would indicate that there is no possible polynomial of this degree that passes through all the points.

2.2 Solving an example nullspace using reduced row echelon form (#35)

Let's find the nullspace for a matrix

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{bmatrix}$$

Remember, $\mathcal{N}(A) = \{x \in \mathbb{R}^n : Ax = 0\}$. So we will look for the vector $x \in \mathbb{R}^4$ below:

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

We will solve this linear system of equations by putting the following matrix into reduced row echelon form:

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

This is essentially the same as putting the unaugmented matrix A into RREF, since there is no elementary row operation that will change the elements in the last column, but either way it gives us

$$\begin{bmatrix} 1 & 0 & -1 & -2 & 0 \\ 0 & 1 & 2 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The first two rows/columns have pivot elements, whose columns are the pivot columns (since they have leading 1's). The rest are free columns. The matrices can be read to be saying

$$x_1 = x_3 + 2x_4 \tag{12}$$
$$x_2 = -2x_3 - 3x_4 \tag{13}$$

Or,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 2 \\ -3 \\ 0 \\ 1 \end{bmatrix}$$

And since, in general, $\mathcal{N}(A) = \mathcal{N}(\mathbf{rref}(A))$, we get

$$\mathcal{N}(A) = \mathbf{span}\left(\begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \\ 0 \\ 1 \end{bmatrix} \right).$$

2.3 Singularity implications

For an n-by-n matrix A

Invertible	mnemonic
$\|A\| \neq 0$	$\|A\| = 0 \Rightarrow$ you can't compute the inverse
non-singular	- (remember base case 2×2 matrix inverse involves $1/\|A\|$ term)
A is full rank	linearly independent columns (invertibility $\Rightarrow$ 1-to-1/injective)
$\mathcal{N}(A) = \{0\}$	linearly independent columns
$\mathcal{R}(A) = \mathbb{R}^n$	linearly independent columns
$Ax = b$ has unique solution for every b	- no more than one solution (can't add members of $\mathcal{N}(A)$ for multiple b) - one solution, since $\mathcal{R}(A) = \mathbb{R}^n$; everything reachable/surjective - one solution found using the unique inverse of A
$\text{rref}(A) = I_n$	
A is a product of elementary matrices	