Final Exam Formulas Sheet - EK 103 Spring 2025

REMINDER!

This sheet (formulas, definitions, and equations) is not a guide for what will be on the exam. You will need to practice, on your own: the procedures for performing calculations, how to approach or solve problems, or how to use the facts on this sheet to interpret your calculations.

1 Markov Chains (Ch. 5.9)

- A matrix $A \in \mathbb{R}^{n \times n}$ is a Markov Matrix, also called a probability matrix or a stochastic matrix or a transition matrix, if all of its columns sum to 1.

  For example:

  $$A = \begin{bmatrix} 0.95 & 0 & 0.03 \\ 0.05 & 0.7 & 0.01 \\ 0 & 0.3 & 0.96 \end{bmatrix}$$

  Confirm:

  $$0.95 + 0.05 + 0 = 1, \qquad 0 + 0.7 + 0.03 = 1, \qquad 0.3 + 0.01 + 0.96 = 1$$

- A dynamical system $\mathbf{x}_{t+1} = A\mathbf{x}_t$ is called a Markov Chain if:
  1. The matrix A is a Markov matrix, and
  2. The vector at time zero, $\mathbf{x}_0$, sums to 1 as well.
- If the A matrix is a Markov matrix, the following is true about its eigenvalues:
  1. $\lambda = 1$ is an eigenvalue.
  2. All remaining eigenvalues are less than 1 in magnitude, $|\lambda| < 1$.
- The steady state solution to a Markov chain is an eigenvector associated with $\lambda = 1$, scaled so that it also sums to 1.
- A state transition diagram has arrows labeled according to the Markov matrix, with the column being "from" and the row being "to."

2 Projections and Least Squares (Ch. 6)

- The projection of a vector $\mathbf{y}$ onto a line defined by the vector $\mathbf{a}$ is both of these two equivalent formulas:

$$\hat{\mathbf{y}} = \frac{\mathbf{y}^\top \mathbf{a}}{\mathbf{a}^\top \mathbf{a}} \mathbf{a} = \frac{\mathbf{a}\mathbf{a}^\top}{\mathbf{a}^\top \mathbf{a}} \mathbf{y}$$

- If you're given a set of vectors $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$, then the projection of a vector $\mathbf{y}$ onto the subspace spanned by these vectors can be found by putting them into a matrix

$$A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots \mathbf{a}_n \end{bmatrix} \in \mathbb{R}^{m \times n}$$

  and performing a different calculation depending on properties of A:

  1. The set of vectors is a basis, or in other words, the columns of A are linearly independent, then:

  $$\hat{\mathbf{y}} = A(A^\top A)^{-1} A^\top \mathbf{y}$$

     The matrix $P = A(A^\top A)^{-1} A^\top$ is called the projection matrix.

  2. If the set of vectors is a basis, and in addition, is orthogonal (each $\mathbf{a}_i^\top \mathbf{a}_j = 0$) and is also normalized (each $||\mathbf{a}_i|| = 1$), then:

  $$\hat{\mathbf{y}} = AA^\top \mathbf{y}$$

- The least squares solution to an inconsistent $A\mathbf{x} = \mathbf{b}$ problem is the vector $\hat{\mathbf{x}}$ found by projecting $\mathbf{b}$ onto the column space of A.

  If the columns of A are linearly independent:

$$\hat{\mathbf{x}} = (A^\top A)^{-1} A^\top \mathbf{b}, \qquad \hat{\mathbf{b}} = A\hat{\mathbf{x}}$$

- If the columns of A are not linearly independent, then the least squares solution is found by parameterizing solutions to the normal equations,

$$A^\top A\mathbf{x} = A^\top \mathbf{b}$$

  ...because $A^\top A$ will not be invertible.

3 Symmetric Matrices and Diagonalization (Ch. 7.1)

1. A matrix P is orthonormal if its columns are orthogonal ($\mathbf{p}_i^\top \mathbf{p}_j = 0$) and normalized ($||\mathbf{p}_i|| = 1$).

2. If a matrix is orthonormal, its inverse equals its transpose: $P^{-1} = P^\top$

3. A matrix A is symmetric if it equals its transpose, $A^\top = A$. In other words, symmetric along its diagonal.

4. If the matrix A is symmetric, then:

   (a) A is always diagonalizable.

   (b) Its eigenvectors are orthogonal.

   (c) It can be orthogonally diagonalized if the matrix P is made orthonormal by choosing normalized eigenvectors:

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

PREVIOUSLY FROM MIDTERM EXAM 2:

4 Matrices, Subspaces, and Inverses: Chapter 2 + 4

- Can I multiply these two matrices? Two matrices can only be multiplied together when the number of columns of the first matrix equals the number of rows of the second matrix. The "inside dimensions" must match.

$$A \in \mathbb{R}^{m\times n}, \qquad B \in \mathbb{R}^{p\times q}, \qquad AB \in \mathbb{R}^{m\times q} \qquad \text{only possible if } n = p$$

- The inverse of a matrix can be calculated by a convenient formula when $A \in \mathbb{R}^{2\times 2}$,

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \qquad A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

  Or more generally, when $A \in \mathbb{R}^{n\times n}$, its inverse can be calculated by row reduction:

$$\begin{bmatrix} A & | & I \end{bmatrix} \qquad \rightarrow \text{Row Reduce} \rightarrow \qquad \begin{bmatrix} I & | & A^{-1} \end{bmatrix}$$

- There are four fundamental subspaces associated with a matrix A.

  1. Column Space: the set of all linear combinations of A's columns.
$$\text{Col } A = \text{span}\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\}$$
  2. Null Space: the set of all solutions to the homogeneous equation, $A\mathbf{x} = \mathbf{0}$.
$$\text{Nul } A = \{\mathbf{x} : A\mathbf{x} = \mathbf{0}\}$$
  3. Row Space: the column space of A transposed.
$$\text{Row } A = \text{Col } A^\top$$
  4. Left Null Space: the null space of A transposed. There isn't a three-letter-symbol for "left null."
$$\text{Nul } A^\top$$

  Note that all of these can be expressed as a linear combination of vectors, i.e., a span! Col and Row are defined that way, and our "parameterizing solutions" procedure does this for Nul.

- A basis for a subspace is a set of linearly independent vectors that spans the subspace.
  ("Not too many, not too few, just right.") For example:

  Valid Basis for $\mathbb{R}^2$
$$\left\{\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}1\\1\end{bmatrix}\right\}, \quad \text{or} \quad \left\{\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}3\\3\end{bmatrix}\right\}$$
  Valid Basis for $\mathbb{R}^3$
$$\left\{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}2\\3\\0\end{bmatrix}, \begin{bmatrix}4\\5\\6\end{bmatrix}\right\}$$

  Not a Valid Basis for $\mathbb{R}^2$
$$\left\{\begin{bmatrix}3\\3\end{bmatrix}, \begin{bmatrix}1\\1\end{bmatrix}\right\}$$
  Not a Valid Basis for $\mathbb{R}^3$
$$\left\{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}2\\3\\0\end{bmatrix}\right\} \quad \text{or} \quad \left\{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}2\\3\\0\end{bmatrix}, \begin{bmatrix}4\\5\\6\end{bmatrix}, \begin{bmatrix}7\\8\\9\end{bmatrix}\right\}$$

- Finding Bases for Spaces:

  THEOREM 6 The pivot columns of a matrix A form a basis for Col A.

  ...and for Nul A, parameterize solutions to $A\mathbf{x} = \mathbf{0}$.

- The dimension of a subspace is the number of linearly independent vectors in a basis.
- The rank of a matrix A is the dimension of its column space, i.e., the number of vectors in a basis for Col A.

$$\dim \text{Col } A = \text{rank } A$$

- The nullity of a matrix A is the dimension of its null space, i.e., the number of vectors in a basis for Nul A.

$$\dim \text{Nul } A = \text{nullity } A$$

- The relationship between the dimension of the column space and the dimension of the null space:

THEOREM 14 **The Rank Theorem**

The dimensions of the column space and the row space of an $m \times n$ matrix A are equal. This common dimension, the rank of A, also equals the number of pivot positions in A and satisfies the equation

$$\text{rank } A + \dim \text{Nul } A = n$$

$$\left\{\begin{array}{c}\text{number of}\\\text{pivot columns}\end{array}\right\} + \left\{\begin{array}{c}\text{number of}\\\text{nonpivot columns}\end{array}\right\} = \left\{\begin{array}{c}\text{number of}\\\text{columns}\end{array}\right\}$$

- The Invertible Matrix Theorem: Many of our properties related to linear independence, existence and uniqueness of solutions, column space, etc., all become the *same* property when a matrix is square.

The Invertible Matrix Theorem

Let A be a square $n \times n$ matrix. Then the following statements are equivalent. That is, for a given A, the statements are either all true or all false.

a. A is an invertible matrix.
b. A is row equivalent to the $n \times n$ identity matrix.
c. A has n pivot positions.
d. The equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.
e. The columns of A form a linearly independent set.
f. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one.
g. The equation $A\mathbf{x} = \mathbf{b}$ has at least one solution for each $\mathbf{b}$ in $\mathbb{R}^n$.
h. The columns of A span $\mathbb{R}^n$.
i. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ maps $\mathbb{R}^n$ onto $\mathbb{R}^n$.
j. There is an $n \times n$ matrix C such that $CA = I$.
k. There is an $n \times n$ matrix D such that $AD = I$.
l. A^T is an invertible matrix.

The Invertible Matrix Theorem (continued)

Let A be an $n \times n$ matrix. Then the following statements are each equivalent to the statement that A is an invertible matrix.

m. The columns of A form a basis of $\mathbb{R}^n$.

n. $\text{Col } A = \mathbb{R}^n$

o. $\dim \text{Col } A = n$

p. $\text{rank } A = n$

q. $\text{Nul } A = \{\mathbf{0}\}$

r. $\dim \text{Nul } A = 0$

The Invertible Matrix Theorem (continued)

Let A be an $n \times n$ matrix. Then A is invertible if and only if:

s. The number 0 is *not* an eigenvalue of A.

t. The determinant of A is *not* zero.

5 Determinants: Chapter 3

- The determinant of a square matrix A is the sum of a "cofactor expansion,"

DEFINITION

For $n \geq 2$, the **determinant** of an $n \times n$ matrix $A = [a_{ij}]$ is the sum of n terms of the form $\pm a_{1j} \det A_{1j}$, with plus and minus signs alternating, where the entries $a_{11}, a_{12}, \ldots, a_{1n}$ are from the first row of A. In symbols,

$$\begin{aligned}\det A &= a_{11} \det A_{11} - a_{12} \det A_{12} + \cdots + (-1)^{1+n} a_{1n} \det A_{1n} \\ &= \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det A_{1j}\end{aligned}$$

...appying the formula iteratively until you get down to a 2x2 matrix,

$$A_{ij} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \qquad \det\, A_{ij} = ad - bc$$

- The determinant of a triangular matrix is the product of entries along its diagonal.
- The determinant of a matrix is equal to the product of its eigenvalues.
  For example, if $A \in \mathbb{R}^{2\times 2}$, then

$$\det(A) = \lambda_1 \cdot \lambda_2$$

- The determinant of a matrix changes when you perform row operations:

THEOREM 3

Row Operations

Let A be a square matrix.

a. If a multiple of one row of A is added to another row to produce a matrix B, then $\det B = \det A$.

b. If two rows of A are interchanged to produce B, then $\det B = -\det A$.

c. If one row of A is multiplied by k to produce B, then $\det B = k \cdot \det A$.

You can use row operations, turning A into its REF'd form U, as an alternative method to calculate the determinant. If you perform r-many row interchanges, and do not scale any rows, then:

$$\det\, A = (-1)^r \det\, U$$

- The determinant calculates the area change between the pre-image (X) and the image ($Y = AX$) of a linear transformation:

$$(\text{area enclosed by } Y) = (\det\, A) \cdot (\text{area enclosed by } X)$$

6 Eigenvalues and Eigenvectors: Chapter 5

- The eigenvalues $\lambda \in \mathbb{R}$ and the eigenvectors $\mathbf{v} \in \mathbb{R}^n$ of a square matrix $A \in \mathbb{R}^{n\times n}$ are the special pairs of scalars/vectors that satisfy the equation:

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

- The characteristic polynomial for a matrix A is $\det(A - \lambda I)$, with λ as an unknown variable.
- The characteristic equation for a matrix A is its characteristic polynomial set equal to zero.

$$\det(A - \lambda I) = 0$$

- The eigenvalues of a matrix can be found by solving its characteristic equation:

> A scalar λ is an eigenvalue of an $n \times n$ matrix A if and only if λ satisfies the characteristic equation
> $$\det(A - \lambda I) = 0$$

- The eigenvectors of a matrix A associated with eigenvalue λ_i can be found by parameterizing solutions to this modified null space problem:

$$(A - \lambda_i I)\mathbf{v}_i = 0 \qquad \rightarrow \qquad \mathbf{v}_i = \ \ldots$$

- The eigenspace of a matrix A associated with its eigenvalue λ_i is the set of all its eigenvectors for λ_i.
- Some facts about eigenvalues and eigenvectors:
  1. Eigenvalues and eigenvectors are only for square matrices. Matrices that are wide or tall do not have eigenvalues or eigenvectors.
  2. An $n \times n$ matrix will always have n eigenvalues, $\{\lambda_1, \ldots, \lambda_n\}$.
  3. But some of those eigenvalues can be repeated (appear more than once).
  4. An $n \times n$ matrix will have **at most** n linearly independent eigenvectors.
- Eigenvalues and independence of eigenvectors:

THEOREM 2 If $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are eigenvectors that correspond to distinct eigenvalues $\lambda_1, \ldots, \lambda_r$ of an $n \times n$ matrix A, then the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ is linearly independent.

- The algebraic multiplicity of eigenvalue λ_i is the number of times it appears in the characteristic equation, i.e., how many times it is repeated.
- The geometric multiplicity of eigenvalue λ_i is how many linearly independent eigenvectors it has, i.e., the dimension of its eigenspace.
- A square matrix $A \in \mathbb{R}^{n\times n}$ is diagonalizable if it can be written as the product of three matrices based on its eigenvalues λ_i and corresponding eigenvectors $\mathbf{v}_i$:

$$A = PDP^{-1}, \qquad D = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \lambda_n \end{bmatrix}, \qquad P = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_1 & \ldots \mathbf{v}_n \end{bmatrix}$$

  Therfore, a matrix is not diagonalizable if P^{-1} does not exist: i.e., if there are "not enough linearly independent eigenvectors."

- A diagonalizable matrix is easy to raise to a matrix power:

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

- Matrices that have unique eigenvalues (all λ_i have an algebraic multiplicty $= 1$) are always diagonalizable.

- Matrices with repeated eigenvalues (at least one λ_i with algebraic multiplicity > 1) are diagonalizable if their geometric multiplicity equals their algebraic multiplicity: i.e., "you get enough eigenvectors in the end."

  For example:

$$A = \text{diagonalizable 3x3 matrix with repeated eigenvalues,}$$

$$\{\lambda_1 = 7,\ \lambda_2 = 6,\ \lambda_3 = 6\}, \qquad \mathbf{v}_1 = x_2 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \qquad \mathbf{v}_2 = x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix},$$

$$\Rightarrow \qquad D = \begin{bmatrix} 7 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{bmatrix}, \qquad P = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

PREVIOUSLY FROM MIDTERM EXAM 1:

7 Solving Systems of Linear Equations, Vector Algebra: Chapter 1 + Chapter 6.1

- A vector can be written as a bold lower-case letter, or a lower-case letter with an arrow above it, both are the same. We write the number of elements in the vector using the symbol $\mathbb{R}$, for example, if there are n elements in the vector:

$$\mathbf{x} = \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$$

You can also be given two points P and Q and write the vector between them as $\overrightarrow{PQ}$ like this:

$$P = (x_1, y_1), \qquad Q = (x_2, y_2), \qquad \overrightarrow{PQ} = \begin{bmatrix} x_2 - x_1 \\ y_2 - y_1 \end{bmatrix}$$

- Transpose: If you have a vector $\mathbf{x}$ that has n components, its transpose is a row:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad \mathbf{x}^\top = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \in \mathbb{R}^{1\times n}$$

- Dot Product: If you have two vectors with the same number of components, their dot product is:

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}, \qquad \mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}, \qquad \mathbf{v}\cdot\mathbf{u} = \mathbf{v}^\top\mathbf{u} = v_1u_1 + v_2u_2 + \ldots + v_nu_n$$

Observe that the dot product is commutative: $\mathbf{u}^\top\mathbf{v}$ is the same calculation as $\mathbf{v}^\top\mathbf{u}$.

- Norm: The norm of a vector is its length, and is calculated by the Pythagorean theorem:

$$||\mathbf{v}|| = \sqrt{\mathbf{v}^\top\mathbf{v}} = \sqrt{v_1^2 + v_2^2 + \ldots + v_n^2}$$

- Dot Product Using Lengths and Angles: Another definition of the dot product uses the lengths of the two vectors, and and angle between them:

$$\mathbf{u}\cdot\mathbf{v} = ||\mathbf{u}||\ ||\mathbf{v}||\ \cos(\theta)$$

- Orthogonality: Two vectors are orthogonal if the angle between them is 90° (or some multiple of 90°, like −90° or 270°). Orthogonality can be determined by a dot product calculation:

> Two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ are **orthogonal** (to each other) if $\mathbf{u}\cdot\mathbf{v} = 0$.

- A system of linear equations occurs when you have a matrix (A) multiplied by a vector of unknowns ($\mathbf{x}$), which must equal another vector ($\mathbf{b}$):

$$A\mathbf{x} = \mathbf{b} \quad \Longleftrightarrow \quad \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \quad \Longleftrightarrow \quad \begin{matrix} a_{11}x_1 + a_{12}x_2 + \ldots a_{1m}x_m = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots a_{2m}x_m = b_2 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots a_{nm}x_m = b_n \end{matrix}$$

- A homogeneous linear system is when the vector $\mathbf{b} = 0$, like $A\mathbf{x} = 0$
- A nonhomogeneous linear system is when the vector $\mathbf{b}$ is not zero, like $A\mathbf{x} = \mathbf{b}$
- An augmented matrix is when we put the $\mathbf{b}$ vector as an additional column, $[A\ \mathbf{b}]$. You can write a vertical line between the A part and the $\mathbf{b}$ part, or not, both are valid.

$$[A \mid \mathbf{b}] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1m} & b_1 \\ a_{21} & a_{22} & \dots & a_{2m} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} & b_n \end{array}\right]$$

- The leading entry of a row in a matrix is its left-most nonzero number.
- The Row Echelon Form (REF) of a matrix is often written as $[U\ \mathbf{c}]$ and the Reduced Row Echelon Form (RREF) is often written as $[R\ \mathbf{d}]$, and are defined as:

DEFINITION

A rectangular matrix is in **echelon form** (or **row echelon form**) if it has the following three properties:

1. All nonzero rows are above any rows of all zeros.
2. Each leading entry of a row is in a column to the right of the leading entry of the row above it.
3. All entries in a column below a leading entry are zeros.

If a matrix in echelon form satisfies the following additional conditions, then it is in **reduced echelon form** (or **reduced row echelon form**):

4. The leading entry in each nonzero row is 1.
5. Each leading 1 is the only nonzero entry in its column.

- A pivot or pivot position is the leading entry in a row when a matrix is in REF, such as the black boxes in this graphic. A pivot column is a column with a pivot in REF.

EXAMPLE 1 The following matrices are in echelon form. The leading entries ($\blacksquare$) may have any nonzero value; the starred entries ($\ast$) may have any value (including zero).

$$
\left[\begin{array}{ccc|c}
\blacksquare & \ast & \ast & \ast \\
0 & \blacksquare & \ast & \ast \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{array}\right],
\quad
\left[\begin{array}{ccccccccc|c}
0 & \blacksquare & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast \\
0 & 0 & 0 & \blacksquare & \ast & \ast & \ast & \ast & \ast & \ast \\
0 & 0 & 0 & 0 & \blacksquare & \ast & \ast & \ast & \ast & \ast \\
0 & 0 & 0 & 0 & 0 & \blacksquare & \ast & \ast & \ast & \ast \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & \ast
\end{array}\right]
\longleftarrow \text{REF}
$$

Basic Free

Basic Free

The following matrices are in reduced echelon form because the leading entries are 1's, and there are 0's below *and above* each leading 1.

$$
\begin{bmatrix}
1 & 0 & \ast & \ast \\
0 & 1 & \ast & \ast \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{bmatrix},
\quad
\begin{bmatrix}
0 & 1 & \ast & 0 & 0 & 0 & \ast & \ast & 0 & \ast \\
0 & 0 & 0 & 1 & 0 & 0 & \ast & \ast & 0 & \ast \\
0 & 0 & 0 & 0 & 1 & 0 & \ast & \ast & 0 & \ast \\
0 & 0 & 0 & 0 & 0 & 1 & \ast & \ast & 0 & \ast \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \ast
\end{bmatrix}
\longleftarrow \text{RREF}
$$

- A basic variable is one of the x_i when the corresponding column contains a pivot in REF.

  In the example on the left, x_1 and x_2 are basic variables.

- A free variable is one of the x_i when the corresponding column does not contain a pivot in REF.

  In the example on the left, x_3 is a free variable.

- A linear combination is a sum of scaled vectors, and is another way to interpret matrix-vector multiplication. Here, the $\mathbf{a}_i$ vectors are the columns of the A matrix:

DEFINITION

If A is an $m \times n$ matrix, with columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$, and if $\mathbf{x}$ is in $\mathbb{R}^n$, then the **product of A and $\mathbf{x}$**, denoted by $A\mathbf{x}$, is **the linear combination of the columns of A using the corresponding entries in $\mathbf{x}$ as weights**; that is,

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

- Existence and Uniqueness: "How many solutions exist to this $A\mathbf{x} = \mathbf{b}$ problem?" There are only three possibilities: either 0, or 1, or ∞ solutions.

  If there is only one solution, we say it is unique. As long as at least one solution exists, we say it is consistent.

THEOREM 2

Existence and Uniqueness Theorem

A linear system is consistent if and only if the rightmost column of the augmented matrix is *not* a pivot column—that is, if and only if an echelon form of the augmented matrix has *no* row of the form

$$[\,0 \quad \cdots \quad 0 \quad b\,] \qquad \text{with } b \text{ nonzero}$$

If a linear system is consistent, then the solution set contains either (i) a unique solution, when there are no free variables, or (ii) infinitely many solutions, when there is at least one free variable.

This definition is equivalent to other facts related to linear combinations, span, and linear independence:

THEOREM 4

Let A be an $m \times n$ matrix. Then the following statements are logically equivalent. That is, for a particular A, either they are all true statements or they are all false.

a. For each $\mathbf{b}$ in $\mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution.

b. Each $\mathbf{b}$ in $\mathbb{R}^m$ is a linear combination of the columns of A.

c. The columns of A span $\mathbb{R}^m$.

d. A has a pivot position in every row.

Be careful when attempting to 'flip' this theorem. For example: if solutions exist to $A\mathbf{x} = \mathbf{b}$ for a particular $\mathbf{b}$ but not others, there must be at least one row without a pivot. HOWEVER! This tells you nothing about *how many columns* have pivots, and does not tell you if solutions are *unique vs. infinite* for that $\mathbf{b}$.

- The trivial solution to a homogeneous linear system, $A\mathbf{x} = \mathbf{0}$, is the vector $\mathbf{x} = 0$. It is always a solution to $A\mathbf{x} = \mathbf{0}$.
- A nontrivial solution to $A\mathbf{x} = \mathbf{0}$ is some other $\mathbf{x}$ that is not 0.

  Fact:

  The homogeneous equation $A\mathbf{x} = \mathbf{0}$ has a nontrivial solution if and only if the equation has at least one free variable.

- The homogeneous solution to an $A\mathbf{x} = \mathbf{b}$ problem is the solution to the $A\mathbf{x} = \mathbf{0}$ problem with the same A matrix.
- The particular solution is a constant vector (*not* multiplied by an unknown) that is a solution to $A\mathbf{x} = \mathbf{b}$.
- The complete solution is the sum of the homogeneous solution and particular solution.

  Fact:

  Suppose the equation $A\mathbf{x} = \mathbf{b}$ is consistent for some given $\mathbf{b}$, and let $\mathbf{p}$ be a solution. Then the solution set of $A\mathbf{x} = \mathbf{b}$ is the set of all vectors of the form $\mathbf{w} = \mathbf{p} + \mathbf{v}_h$, where $\mathbf{v}_h$ is any solution of the homogeneous equation $A\mathbf{x} = \mathbf{0}$.

- A set of vectors is linearly independent if, when put as an $A\mathbf{x} = \mathbf{0}$ problem, the only solution is $\mathbf{x} = 0$. (the trivial solution!)

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \ldots + x_n\mathbf{a}_n = 0$$

If $\mathbf{x} = 0$ is the only solution, then the vectors $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\}$ are linearly independent.

- A set of vectors is linearly dependent if it is not linearly independent. In other words, dependent if nontrivial solutions to $A\mathbf{x} = \mathbf{0}$.

- Linear Independence and Row Reduction: The linear independence of a set of vectors $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\}$ can be determined by putting them as columns of a matrix $A = [\mathbf{a}_1\ \mathbf{a}_2\ \ldots \mathbf{a}_n]$ and calculating the REF of that matrix:

> The columns of a matrix A are linearly independent if and only if the equation $A\mathbf{x} = \mathbf{0}$ has *only* the trivial solution. (3)

This implies two additional facts:

THEOREM 8

> If a set contains more vectors than there are entries in each vector, then the set is linearly dependent. That is, any set $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ in $\mathbb{R}^n$ is linearly dependent if $p > n$.

$$\begin{array}{cc} & p \\ n & \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix} \end{array}$$

...since at least one column cannot have a pivot in it.

THEOREM 9

> If a set $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ in $\mathbb{R}^n$ contains the zero vector, then the set is linearly dependent.

...since a column of zeros cannot have a pivot in it.

- The span of a set of vectors is all linear combinations of them:

DEFINITION

> If $\mathbf{v}_1, \ldots, \mathbf{v}_p$ are in $\mathbb{R}^n$, then the set of all linear combinations of $\mathbf{v}_1, \ldots, \mathbf{v}_p$ is denoted by Span $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ and is called the **subset of** $\mathbb{R}^n$ **spanned** (or **generated**) **by** $\mathbf{v}_1, \ldots, \mathbf{v}_p$. That is, Span $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ is the collection of all vectors that can be written in the form
>
> $$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p$$
>
> with $c_1, \ldots, c_p$ scalars.

Asking whether a vector $\mathbf{b}$ is in Span $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ amounts to asking whether the vector equation

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_p\mathbf{v}_p = \mathbf{b}$$

has a solution, or, equivalently, asking whether the linear system with augmented matrix $[\,\mathbf{v}_1\ \cdots\ \mathbf{v}_p\ \ \mathbf{b}\,]$ has a solution.