ROB 101 - Fall 2021

The Vector Space $\mathbb{R}^n$: Part 1

October 4, 2021

UNIVERSITY OF MICHIGAN

Learning Objectives

- Instead of working with individual vectors, we will work with a collection of vectors.
- Our first encounter with some of the essential concepts in Linear Algebra that go beyond systems of equations.

- Vectors as n-tuples of real numbers
- $\mathbb{R}^n$ as the collection of all n-tuples of real numbers
- Linear combinations of vectors
- Linear independence of vectors
- Relation of these concepts to the existence and uniqueness of solutions to $Ax = b$.
- LU Factorization to check the linear independence of a set of vectors, and LDLT to check if one vector is a linear combination of a set of vectors.

Recall: Concise Definition of Linear Independence

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_m v_m = 0_{n\times 1} \iff \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = 0_{m\times 1}.$$

Recall: Pro Tip for Checking Linear Independence

Corollary

$$\boxed{A\alpha = 0 \iff \left(A^{\mathsf{T}} A\right)\alpha = 0.}$$

Recall: Pro Tip for Checking Linear Independence

Consider the vectors in $\mathbb{R}^n$,

$$\left\{ v_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix}, v_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix}, ..., v_m = \begin{bmatrix} a_{1m} \\ a_{2m} \\ \vdots \\ a_{nm} \end{bmatrix} \right\},$$

and use them as the columns of a matrix that we call A,

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}.$$

Recall: Pro Tip for Checking Linear Independence

The following statements are equivalent:

- The set of vectors $\{v_1, v_2, \ldots, v_m\}$ is linearly independent.
- The $m \times m$ matrix $A^\mathsf{T} \cdot A$ is invertible.
- $\det(A^\mathsf{T} \cdot A) \neq 0$.
- For any LU Factorization $P \cdot (A^\mathsf{T} \cdot A) = L \cdot U$ of $A^\mathsf{T} A$, the $m \times m$ upper triangular matrix U has no zeros on its diagonal.

Example

Determine if the vectors

$$v_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix}, v_2 = \begin{bmatrix} 0 \\ -2 \\ 3 \\ 1 \end{bmatrix}, v_3 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}$$

are linearly independent or dependent.

We form the matrix

$$A := \left[\begin{array}{ccc} v_1 & v_2 & v_3 \end{array}\right] = \left[\begin{array}{rrr} 1 & 0 & 1 \\ 2 & -2 & 0 \\ 0 & 3 & -1 \\ 1 & 1 & 0 \end{array}\right].$$

We form the matrix

$$A := \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 2 & -2 & 0 \\ 0 & 3 & -1 \\ 1 & 1 & 0 \end{bmatrix}.$$

We check that $\det(A^{\mathsf{T}}A) = 100$. We conclude the set of vectors is linearly independent.

Remark

We have $A^\mathsf{T}A = \begin{bmatrix} 6 & -3 & 1 \\ -3 & 14 & -3 \\ 1 & -3 & 2 \end{bmatrix}$. *Notice the symmetry in this matrix. We already know that*

$$(A^\mathsf{T}A)^\mathsf{T} = A^\mathsf{T}(A^\mathsf{T})^\mathsf{T} = A^\mathsf{T}A.$$

Remark

We have $A^{\mathsf{T}}A = \begin{bmatrix} 6 & -3 & 1 \\ -3 & 14 & -3 \\ 1 & -3 & 2 \end{bmatrix}$. *Notice the symmetry in this matrix. We already know that*

$$(A^{\mathsf{T}}A)^{\mathsf{T}} = A^{\mathsf{T}}(A^{\mathsf{T}})^{\mathsf{T}} = A^{\mathsf{T}}A.$$

A square matrix B that is equal to its transpose is called a symmetric matrix.

$$\boxed{B \text{ is symmetric} \iff B = B^{\mathsf{T}}}$$

Existence and Uniqueness of Solutions to $Ax = b$

The following two statements are equivalent

1. The system of linear equations $Ax = b$ has a solution, and, it is unique.

2. b is a linear combination of the columns of A, and, the columns of A are linearly independent.

Existence and Uniqueness of Solutions to $Ax = b$

Remark

If b can be expressed as a linear combination of the columns of A, then the matrix equation $Ax = b$ is consistent and has a solution.

If the columns of A are linearly independent, then $\alpha = 0$ is the only solution of $A\alpha = 0$ and the solution is unique.

Does the following rectangular system have a solution? If it does, is it unique?

$$A = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 2 & -2 & 0 \\ 0 & 3 & -1 \\ 1 & 1 & 0 \end{bmatrix} \quad \text{and} \quad b = \begin{bmatrix} 3 \\ 4 \\ -5 \\ 0 \end{bmatrix}.$$

Uniqueness test: We check that $\det(A^{\mathsf{T}}A) = 100$. We conclude the set of vectors is linearly independent. Hence, if there is a solution, then the solution is unique.

Existence test: To check for the existence of the solution we need to check if b is a linear combination of the columns of A. We can construct the following matrix

$$A_e := \begin{bmatrix} A & b \end{bmatrix} = \{v_1, v_2, v_3, b\}.$$

Existence test: To check for the existence of the solution we need to check if b is a linear combination of the columns of A. We can construct the following matrix

$$A_e := \begin{bmatrix} A & b \end{bmatrix} = \{v_1, v_2, v_3, b\}.$$

Now clearly if b is a linear combination of the columns of A, then the sets $\{v_1, v_2, v_3\}$ and $\{v_1, v_2, v_3, b\}$ must have the same *number of linearly independent vectors* (3 in this example).

Number of Linearly Independent Vectors in a Set

Consider a finite set of vectors in $\mathbb{R}^n$, such as $\{v_1, \ldots, v_m\}$.

Q1. Can we define the maximum number if linearly independent vectors in the set?

Number of Linearly Independent Vectors in a Set

Consider a finite set of vectors in $\mathbb{R}^n$, such as $\{v_1, \ldots, v_m\}$.

Q1. Can we define the maximum number if linearly independent vectors in the set?

Q2. Why this question is interesting (important)?

Number of Linearly Independent Vectors in a Set

- When does $Ax = b$ have a solution?

- When does $Ax = b$ have a solution?

1. b is a linear combination of the columns of A.

Number of Linearly Independent Vectors in a Set

- When does $Ax = b$ have a solution?

1. b is a linear combination of the columns of A.
2. b is not linearly independent from the columns of A.

Number of Linearly Independent Vectors in a Set

- When does $Ax = b$ have a solution?

1. b is a linear combination of the columns of A.
2. b is not linearly independent from the columns of A.
3. The sets $\{v_1, \ldots, v_m\}$ and $\{v_1, \ldots, v_m, b\}$ must have the same number of linearly independent vectors!

Uber Pro-Tip: Number of Linearly Independent Vectors via an Enhanced LU Factorization

Consider the vectors in $\mathbb{R}^n$,

$$\left\{ v_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix}, v_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix}, ..., v_m = \begin{bmatrix} a_{1m} \\ a_{2m} \\ \vdots \\ a_{nm} \end{bmatrix} \right\},$$

and use them as the columns of a matrix that we call A,

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}.$$

Uber Pro-Tip: Number of Linearly Independent Vectors via an Enhanced LU Factorization

Fact

The matrix $A^\mathsf{T} \cdot A$ always has an LDLT Factorization

$$P \cdot A^\mathsf{T} \cdot A \cdot P^\mathsf{T} = L \cdot D \cdot L^\mathsf{T},$$

Uber Pro-Tip: Number of Linearly Independent Vectors via an Enhanced LU Factorization

Fact

The matrix $A^\mathsf{T} \cdot A$ always has an LDLT Factorization

$$P \cdot A^\mathsf{T} \cdot A \cdot P^\mathsf{T} = L \cdot D \cdot L^\mathsf{T},$$

where

- P *is a (row) permutation matrix;*

Uber Pro-Tip: Number of Linearly Independent Vectors via an Enhanced LU Factorization

Fact

The matrix $A^\mathsf{T} \cdot A$ always has an LDLT Factorization

$$P \cdot A^\mathsf{T} \cdot A \cdot P^\mathsf{T} = L \cdot D \cdot L^\mathsf{T},$$

where

- P *is a (row) permutation matrix;*
- P^T*, the transpose of* P*, permutes the columns of* $A^\mathsf{T} A$*;*

Uber Pro-Tip: Number of Linearly Independent Vectors via an Enhanced LU Factorization

Fact

The matrix $A^\mathsf{T} \cdot A$ always has an LDLT Factorization

$$P \cdot A^\mathsf{T} \cdot A \cdot P^\mathsf{T} = L \cdot D \cdot L^\mathsf{T},$$

where

- P *is a (row) permutation matrix;*
- P^T*, the transpose of* P*, permutes the columns of* $A^\mathsf{T} A$*;*
- L *is uni-lower triangular and* L^T*, the transpose of* L*, is therefore uni-upper triangular; and*

Uber Pro-Tip: Number of Linearly Independent Vectors via an Enhanced LU Factorization

Fact

The matrix $A^\mathsf{T} \cdot A$ always has an LDLT Factorization

$$P \cdot A^\mathsf{T} \cdot A \cdot P^\mathsf{T} = L \cdot D \cdot L^\mathsf{T},$$

where

- P *is a (row) permutation matrix;*
- P^T*, the transpose of* P*, permutes the columns of* $A^\mathsf{T}A$*;*
- L *is uni-lower triangular and* L^T*, the transpose of* L*, is therefore uni-upper triangular; and*
- D *is diagonal and has non-negative entries.*

Uber Pro-Tip: Number of Linearly Independent Vectors via an Enhanced LU Factorization

Fact

The matrix $A^\mathsf{T} \cdot A$ *always has an LDLT Factorization*

$$P \cdot A^\mathsf{T} \cdot A \cdot P^\mathsf{T} = L \cdot D \cdot L^\mathsf{T}.$$

Moreover,

- *the number of linearly independent columns of* A *is equal to the number of non-zero entries on the diagonal of* D.

Remark

The LDLT factorization (also called a Cholesky factorization) may look intimidating, but once you realize that $U := D \cdot L^{\mathsf{T}}$ *is upper triangular, this is really a refined LU Factorization that is possible for matrices of the form* $A^{\mathsf{T}} \cdot A$.

Remark

We check that

$$\begin{aligned}\left(P \cdot A^{\mathsf{T}} \cdot A \cdot P^{\mathsf{T}}\right)^{\mathsf{T}} &= \left(P^{\mathsf{T}}\right)^{\mathsf{T}} \cdot (A)^{\mathsf{T}} \cdot \left(A^{\mathsf{T}}\right)^{\mathsf{T}} \cdot (P)^{\mathsf{T}} \\ &= P \cdot A^{\mathsf{T}} \cdot A \cdot P^{\mathsf{T}}.\end{aligned}$$

$P \cdot A^\mathsf{T} \cdot A \cdot P^\mathsf{T}$ is Symmetric

Remark

We check that

$$\begin{aligned}\left(P \cdot A^\mathsf{T} \cdot A \cdot P^\mathsf{T}\right)^\mathsf{T} &= \left(P^\mathsf{T}\right)^\mathsf{T} \cdot (A)^\mathsf{T} \cdot \left(A^\mathsf{T}\right)^\mathsf{T} \cdot (P)^\mathsf{T} \\ &= P \cdot A^\mathsf{T} \cdot A \cdot P^\mathsf{T}.\end{aligned}$$

Without the P^T on the right, the term $P \cdot A^\mathsf{T} \cdot A$ alone would not be symmetric in general. A similar computation shows that $L \cdot D \cdot L^\mathsf{T}$ is also symmetric.

Remark

LDLT (Cholesky) factorization is a particular case of the LU because it is only applicable to symmetric matrices of the form $A^{\mathsf{T}}A$*. However, it is twice as fast as LU!*

Remark

LDLT (Cholesky) factorization is a particular case of the LU because it is only applicable to symmetric matrices of the form $A^\mathsf{T} A$*. However, it is twice as fast as LU!*

It is not a coincidence that we preserve the symmetry of the original matrix during the factorization ($A^\mathsf{T} A \to LDL^\mathsf{T}$)*, and we gain substantial computational savings. Symmetries exist in many natural (and mathematical) problems and preserving such structures often leads to certain advantages in practice.*

Time Your Factorization!

Problem

In Julia, generate symmetric matrices of the form $A^\mathsf{T} A$ by increasing the size of $A^\mathsf{T} A$ by a factor, e.g., 1, 10, 100, 1000, 10000, 100000. Then compute LU and LDLT and record the elapsed time for each algorithm. Create a table where the rows are recorded times, and the columns are for LU and LDLT. Plot your points to compare their performance. In the next chapter, you will learn how to fit a line to these points!

Example

Consider $Ax = b$, where

$$A = \begin{bmatrix} -0.2 & -0.2 & -0.4 & 0.3 & 0.3 \\ 0.3 & 1.0 & -0.1 & -1.1 & -1.7 \\ 0.7 & -1.9 & 1.5 & -0.0 & -3.0 \\ 0.9 & -1.0 & -0.7 & 0.6 & -1.8 \\ -0.5 & 0.8 & -1.1 & -0.5 & -0.5 \\ -2.0 & -0.9 & -0.5 & 0.2 & 0.3 \\ -1.0 & 0.6 & 0.7 & -0.9 & 0.2 \end{bmatrix}_{7\times 5} \quad \text{and } b = \begin{bmatrix} -0.5 \\ 0.1 \\ 0.3 \\ -0.2 \\ -1.3 \\ -3.2 \\ -0.6 \end{bmatrix}_{7\times 1}$$

Does it have a solution? If it does, is it unique?

We form $A^\mathsf{T} A$ and compute in Julia its LDLT Factorization and report the diagonal of D as a row vector

$$\mathrm{diag}(D) = \begin{bmatrix} 15.6 & 5.2 & 4.4 & 2.3 & 0.0 \end{bmatrix}_{1\times 5}$$

We form $A^{\mathsf{T}}A$ and compute in Julia its LDLT Factorization and report the diagonal of D as a row vector

$$\mathrm{diag}(D) = \begin{bmatrix} 15.6 & 5.2 & 4.4 & 2.3 & 0.0 \end{bmatrix}_{1\times 5}$$

- There is a single zero on the diagonal and four non-zero elements. Thus we know that exactly four of the five columns of A are linearly independent.

Example

We form $A^{\mathsf{T}}A$ and compute in Julia its LDLT Factorization and report the diagonal of D as a row vector

$$\mathrm{diag}(D) = \left[\begin{array}{ccccc} 15.6 & 5.2 & 4.4 & 2.3 & 0.0 \end{array}\right]_{1\times 5}$$

- There is a single zero on the diagonal and four non-zero elements. Thus we know that exactly four of the five columns of A are linearly independent.
- Hence, if there does exist a solution to $Ax = b$, it will not be unique.

Next, we form $A_e := [A \;\; b]$ and compute in Julia the LDLT Factorization of $A_e^\mathsf{T} A_e$. We report the diagonal of D_e written as a row vector

$$\mathrm{diag}(D_e) = \left[\begin{array}{cccccc} 15.6 & 12.6 & 3.6 & 2.6 & 0.0 & 0.0 \end{array}\right]_{1\times 6}.$$

Next, we form $A_e := [A \;\; b]$ and compute in Julia the LDLT Factorization of $A_e^{\mathsf{T}} A_e$. We report the diagonal of D_e written as a row vector

$$\operatorname{diag}(D_e) = \left[\begin{array}{cccccc} 15.6 & 12.6 & 3.6 & 2.6 & 0.0 & 0.0 \end{array}\right]_{1\times 6}.$$

- We note that it also has four non-zero entries. We deduce that b is a linear combination of the columns of A.

Next, we form $A_e := [A \;\; b]$ and compute in Julia the LDLT Factorization of $A_e^\mathsf{T} A_e$. We report the diagonal of D_e written as a row vector

$$\mathrm{diag}(D_e) = \left[\begin{array}{cccccc} 15.6 & 12.6 & 3.6 & 2.6 & 0.0 & 0.0 \end{array}\right]_{1\times 6}.$$

- We note that it also has four non-zero entries. We deduce that b is a linear combination of the columns of A.
- Hence, the system of linear equations has a solution and it is not unique. Therefore, it has an infinite number of solutions!

We select a different right hand side and report the key result of its LDLT Factorization,

$$\tilde{b} = \begin{bmatrix} 0.6 \\ 0.4 \\ 0.9 \\ 1.0 \\ 0.8 \\ 0.8 \\ 0.9 \end{bmatrix}$$

$$\Longrightarrow \ \mathrm{diag}(\tilde{D}_e) = \begin{bmatrix} 15.6 & 5.2 & 1.2 & 2.3 & 0.9 & 0.0 \end{bmatrix}_{1\times 6}.$$

Example

$$\mathrm{diag}(\tilde{D}_e) = \begin{bmatrix} 15.6 & 5.2 & 1.2 & 2.3 & 0.9 & 0.0 \end{bmatrix}_{1\times 6}.$$

- This time $\mathrm{diag}(\tilde{D}_e)$ has five non-zero entries, whereas $\mathrm{diag}(D)$ had four non-zero entries.

$$\mathrm{diag}(\tilde{D}_e) = \left[\begin{array}{cccccc} 15.6 & 5.2 & 1.2 & 2.3 & 0.9 & 0.0 \end{array} \right]_{1\times 6}.$$

- This time $\mathrm{diag}(\tilde{D}_e)$ has five non-zero entries, whereas $\mathrm{diag}(D)$ had four non-zero entries.
- Hence, $\tilde{b}$ is not a linear combination of the columns of A, and the system of equations, with this new right hand side, does not have a solution.

Next Time

- Euclidean Norm, Least Squared Error Solutions to Linear Equations, and Linear Regression
- Read Chapter 8 of ROB 101 Book