15.455x Mathematics Methods of Quantitative Finance MIT Sloan School of Management

Review of Linear Algebra

Paul F. Mende

About the word "solution"... a warm-up exercise

• How many solutions to the following? Write down your answers now.

$$3x = 6$$

$$3x = y$$

$$x^{3} = 1$$

$$x^{2} + y^{2} = 1$$

$$\begin{cases} x + 2y = 8 \\ 3x + 4y = 6 \end{cases}$$

$$\begin{cases} x + 2y + 3z + 4t = 0 \\ 5x + 6y + 7z + 8t = 0 \\ x + 2y + 4z + 8t = 2 \end{cases}$$

$$(5)$$

Vectors and vector spaces

• A **vector space** consists of a set of elements, called **vectors**, that is **closed** under the operations of **vector addition** and **scalar multiplication**.

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} + 2 \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 5 \\ 6 \end{pmatrix}$$

Properties of vector addition

- Commutative
- Associative
- Identity
- Inverse

$$\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$$

$$(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$$

$$\mathbf{v} + \mathbf{0} = \mathbf{v}$$

$$\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$$

Properties of scalar multiplication

Scalars form a field (here, the real numbers)

• Associative
$$a(b\mathbf{v}) = (ab)\mathbf{v}$$

• Distributive
$$a(\mathbf{v} + \mathbf{w}) = a\mathbf{v} + a\mathbf{w} \\ (a+b)\mathbf{v} = a\mathbf{v} + b\mathbf{v}$$

• Identity
$$1 \cdot \mathbf{v} = \mathbf{v}$$

Buzzword bingo

- Definitions
 - ▶ Vectors
 - ▶ Vector space
 - ▶ Addition
 - ▶ Scalar multiplication
 - ▶ Closure
 - ▶ Subspace
 - ▶ Dimension
 - ▶Span
 - ▶ Basis
 - ▶ Kernel
 - ▶Image

- ▶ Rank
- ▶ Nullity
- Null space
- → Singular matrix
- ▶ Linear transformation
- ▶ Linear operator
- ▶ Dual space
- ▶ Eigenvalue
- ▶ Eigenvector
- ▶ Diagonalization
- ▶ Change of basis
- ▶ Adjoint transformation

- Inner product
- → Orthogonal
- ▶ Quadratic form
- ▶ Gram-Schmidt process
- **)**

Linear dependence, basis, and dimension

Linear dependence

• A linear combination is a sum of vectors multiplied by arbitrary scalars

$$\mathbf{w} = \sum_{i} a_{i} \mathbf{v_{i}}$$

• A set of vectors is linearly dependent if there are constants, not all zero, such that

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \dots + a_n\mathbf{v}_n = 0$$

• If no such set of constants exist, the set of vectors is linearly independent.

Linear dependence

 Linear dependence means we can write at least one of the vectors in terms of the others.

$$\mathbf{v}_1 = -\frac{1}{a_1}(a_2\mathbf{v}_2 + \dots + a_n\mathbf{v}_n)$$

(Here is where we use that the scalars must be a **field** so that there is an inverse for non-zero scalars.)

 So is there a finite set of vectors that be used to express all the others as linear combinations?

Spanning sets

• One way is to **define** a vector space so that it's true. We call the set of all linear combinations of a given set of vectors the **span** of that set.

$$\operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\} = \left\{ \sum_{i=1}^k a_i \mathbf{v}_i, \quad \forall a_i \in \mathbb{R} \right\}$$

- Because it is to designed to be closed under addition and scalar multiplication, it forms a vector space.
- If every element of a vector space *V* can be expressed as a linear combination of a given set, then that set is said to **span the vector space**.

Basis

- If, in addition to spanning *V*, the vectors in the spanning set are linearly independent, then they form a **basis** for *V*.
 - A basis is a minimal, independent set of vectors that spans the space.
 - The number of vectors in the basis set is called the **dimension** of the vector space.
 - The choice of basis vectors is **not unique**.
 - ▶ Changing the basis, however, does not change the dimension.

Coordinates and notation

• Given a basis, the expression of any vector as a linear combination in terms of the basis vectors is **unique**.

$$\mathbf{v} = c_1 \mathbf{u_1} + c_2 \mathbf{u_2} + \dots + c_n \mathbf{u_n}$$

- The coefficients are called the coordinates with respect to the basis.
- We use vector notation to denote this linear combination compactly as

$$\mathbf{v} = c_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \dots c_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

Subspaces

- A **subspace** S of a vector space V is a vector space which is a subset of V.
- The span of a **subset** of basis vectors of *V* defines a subspace of *V*.
- Any linearly independent set of k < n vectors of V defines a subspace.
 - The set can constitute a basis of the subspace.
 - ▶ The dimension of the subspace is *k*.
 - \blacktriangleright Example: polynomials of degree 3 or less vanishing at x=1.

Linear transformations

- Functions and mappings
- Linear functions
- A linear transformation on a vector space $T: V \to W$ obeys

$$T(\mathbf{v}_1 + \mathbf{v}_2) = T\mathbf{v}_1 + T\mathbf{v}_2,$$

 $T(c\mathbf{v}) = cT\mathbf{v}.$

• This simple property of **linearity** means that any linear transformation is completely described by its action on the basis vectors of a space.

Linear transformations

• Consider the transformation *T* acting on an arbitrary vector, which is expressed as a linear combination of basis vectors. Then by linearity,

$$T\mathbf{v} = T (c_1\mathbf{u_1} + c_2\mathbf{u_2} + \dots + c_n\mathbf{u_n})$$

= $c_1(T\mathbf{u_1}) + c_2(T\mathbf{u_2}) + \dots + c_n(T\mathbf{u_n})$

• Therefore if we know how *T* acts on **each** basis vector in the vector space *V*, we can express the action of *T* on **any** vector by taking linear combinations of these *n* results.

Matrix of a linear transformation

- Let's use column notation for vectors in the target space W.
- T's action on each basis vector of V gives some vector in W, so let's write them in general form as

$$T\mathbf{u_1} = \begin{pmatrix} m_{11} \\ m_{21} \\ \vdots \\ m_{s1} \end{pmatrix}, \quad T\mathbf{u_2} = \begin{pmatrix} m_{12} \\ m_{22} \\ \vdots \\ m_{s2} \end{pmatrix}, \cdots \quad T\mathbf{u_n} = \begin{pmatrix} m_{1n} \\ m_{2n} \\ \vdots \\ m_{sn} \end{pmatrix}$$

• *T* is then characterized by *n* column vectors (the dimension of *V*), each of length *s* (the dimension of *W*).

Matrix of a linear transformation

• Combine these *n* columns to form the matrix *M* corresponding to the linear transformation.

$$M = \begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1n} \\ m_{21} & m_{22} & \cdots & m_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ m_{s1} & m_{s2} & \cdots & m_{sn} \end{pmatrix}$$

- The matrix *M* depends on the choice of bases in *V* and *W*.
- When *M* acts on a column vector of *V*, the result will be a linear combination of the columns of *M*.

Matrix of a linear transformation

• In column notation,

$$M\mathbf{x} = \begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1n} \\ m_{21} & m_{22} & \cdots & m_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ m_{s1} & m_{s2} & \cdots & m_{sn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} m_{11}x_1 + m_{12}x_2 + \cdots + m_{1n}x_n \\ m_{21}x_1 + m_{22}x_2 + \cdots + m_{2n}x_n \\ \vdots \\ m_{s1}x_1 + m_{s2}x_2 + \cdots + m_{sn}x_n \end{pmatrix}$$

In components, this transformation rule reads

$$(M\mathbf{x})_i = \sum_{j=1}^n m_{ij} x_j, \qquad i = 1, 2, \dots, s$$

Linear transformations of the plane

• In two dimensions, let's write this as

$$M\mathbf{x} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}$$

 Acting on this equation from the left with a new linear transformation gives a rule for multiplication of matrices

$$M'M = \begin{pmatrix} a'a + b'c & a'b + b'd \\ c'a + d'c & c'b + d'd \end{pmatrix}$$

• Each column of the product is the result of acting with M' on the corresponding column of M.

Matrix multiplication

· Matrix multiplication defined by composition of linear transformations

$$(M'M)_{ij} = \sum_{k=1}^{s} M'_{ik} M_{kj}$$

- Properties:
 - ▶ Associative
 - ▶ Distributive
 - ▶ NOT commutative
 - ▶ Identity

$$M_1(M_2M_3) = (M_1M_2)M_3$$

$$M_1(M_2 + M_3) = M_1 M_2 + M_1 M_3$$

$$M_1M_2 \neq M_2M_1$$
 (in general)

$$MI = IM = M$$

• (Bonus fact: Matrices also form a vector space of their own.)

Image and kernel

• Two important subspaces can be defined with respect to a linear operator

$$T:V\to W$$

• The **image** of *T* is the set of all vectors in *W* that can be reached from *V*

$$\operatorname{Im} T = \{ \mathbf{w} | \exists \mathbf{v} \in V, \quad T\mathbf{v} = \mathbf{w} \} \subset W$$

The kernel of T is the set of all vectors in V "annihilated" by T

$$\operatorname{Ker} T = \{ \mathbf{v} \in V | T\mathbf{v} = 0 \} \subset V$$
 arbitrage: get nothing from something

Image and kernel

- The **rank** of a linear transformation is the dimension of the image. It is the number of linearly independent columns of a matrix.
- Fundamental Theorem of Linear Transformations:

$$\dim V = \dim(\operatorname{Im} T) + \dim(\operatorname{Ker} T)$$

- If the kernel is empty, i.e., has dimension=0, then the rank of the transformation is the dimension of *V*.
- If in addition, *V* and *W* have the same dimension, then *T* is **invertible**.

Some properties of determinants

- A square matrix has an **inverse** if and only if Det *M* is non-zero
- For a 2x2 matrix, $\operatorname{Det} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad bc$
- Product rule: Det(MM') = (Det M)(Det M')
- Scalar multiplication: $\operatorname{Det}(cM) = c^n \operatorname{Det} M, \quad M: \mathbb{R}^n \to \mathbb{R}^n$
- A singular matrix, where Det M=0, has a non-trivial kernel

Some properties of determinants

- Det *M* is **linear** as function of its individual rows or columns
- Det M is antisymmetric under interchange of adjacent rows or columns
- If any rows or columns are linearly dependent, Det M = 0.
- The determinant is the (oriented) volume of the image of the unit (hyper)cube

Some properties of the trace

- The trace is defined as the sum of the diagonal elements of a square matrix.
- The trace is **invariant** under a change of basis.
- The trace of the **identity** matrix in an *n*-dimensional space is *n*.
- The trace of a product is invariant under cyclic changes in the order; e.g.,

$$\operatorname{Tr} AB = \operatorname{Tr} BA$$
$$\operatorname{Tr} ABC = \operatorname{Tr} BCA = \operatorname{Tr} CAB$$

Matrix inverse

• For 2x2 matrices, the inverse when Det M is non-zero is given by

$$M^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

• The inverse acts on the left or the right to give the identity matrix

$$M(M^{-1}) = (M^{-1})M = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Matrices and inverses

- Example: Rotation matrices
 - ▶ Columns are **image** of basis vectors
 - ▶ Single parameter for angle
 - No fixed points
 - ▶ Determinant = 1
 - ▶ Inverse matrix = reverse rotation

$$M_{\theta} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}, \quad M_{-\theta} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}$$

- For matrices larger than 2X2, use a computer
- Distinguish functions of linear algebra from other operators
 - ► Examples (RTM!)
 - ◆A %*% B Matrix multiplication of A and B
 - ◆A %*% A Square of a matrix
 - ◆A * B Element-by-element multiplication of components
 - +exp(M) Exponential of elements, not exponential of the matrix
- · Beware of numerical issues and instabilities
 - ▶ Reals, rounding
 - ▶ Zero

$$x + 2y + 3z = 3$$
$$4x + 5y + 6z = 6$$
$$7x + 8y + 9z = 9$$

• Write in matrix form $M\mathbf{v}=\mathbf{b}$ and consider solution of form $\mathbf{v}=M^{-1}\mathbf{b}$

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \\ 9 \end{pmatrix}$$

```
> M <- matrix(c(1,2,3,4,5,6,7,8,9),byrow=TRUE,nrow=3)
> M
     [,1] [,2] [,3]
[1,]
[2,]
[3,]
                   an inverse doesn't exist
> solve(M)
Error in solve.default(M) :
  system is computationally singular: reciprocal condition number = 1.54198e-18
             linearly dependent: Notice that the middle column is the average of the first and the third column
> det(M)
[1] 6.661338e-16
> sum(diag(M))
[1] 15
```

• Example: solve

$$x + 2y + 3z = 3$$
$$4x + 5y + 6z = 6$$
$$7x + 8y + 10z = 9$$

• Write in matrix form $M{f v}={f b}$ and consider solution of form ${f v}=M^{-1}{f b}$

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \\ 9 \end{pmatrix}$$

```
> M <- matrix(c(1,2,3,4,5,6,7,8,10),byrow=TRUE,nrow=3)</pre>
> b < - c(3,6,9)
> det(M)
[1] -3
> Minv <- solve(M) ; Minv</pre>
           [,1]
                 [,2] [,3]
[1,] -0.6666667 -1.333333
[2,] -0.6666667 3.666667
[3,] 1.0000000 -2.000000
> v <- Minv %*% b ; v
              [,1]
[1,] -1.000000e+00
[2,] 2.000000e+00
[3,] 1.776357e-15
> M %*% V
     [,1]
[1,]
        6
[2,]
[3,]
        9
```

Systems of linear equations

Systems of linear equations

Let's consider two kinds of systems of linear equations where in general we have s
 equations with n unknowns.

$$\begin{cases} M\mathbf{v} = \mathbf{b} & \text{inhomogeneous, or} \\ M\mathbf{v} = 0 & \text{homogeneous,} \end{cases}$$
 where $\mathbf{v} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^s$, $M : \mathbb{R}^n \to \mathbb{R}^s$

- Expect to find, roughly, that the "inhomogeneous" equation has
 - **▶One** solution if s = n
 - **No** solutions if s > n
 - ▶ Infinitely many solutions if s < n
- The exact situation depends on the dimension of the image and kernel...
 - If r = rank(M) is smaller than it could be, nature of solutions changes.

Case 1: s = n

- When *s*=*n*, there are **the same number** of equations as unknowns, and the matrix *M* is square. As we have seen, there are two sub-cases:
- If det $M \neq 0$, then M is invertible and there is a **unique solution**,

$$M\mathbf{v} = \mathbf{b} \implies \mathbf{v} = M^{-1}\mathbf{b}$$

• If det M=0, then the m equations are **not independent**. M has a **non-zero kernel**, and the rank is less than the dimension of the target space.

$$\dim(\operatorname{Im} M) = r = n - \dim(\ker M) < n = s$$

• So there will be some vectors **b** for which there is no solution.

Case 1: s = n and non-singular

• R example:

$$x + 2y + 3z = 1$$
$$4x + 5y + 6z = 2$$
$$7x + 8y + 10z = 3$$

- ▶ solve(M) inverts matrix
- ▶ solve(M,b) gives unique solution

Solution:

$$\mathbf{v} = \begin{pmatrix} -1/3 \\ 2/3 \\ 0 \end{pmatrix}$$

```
> M <- matrix(c(1,2,3,4,5,6,7,8,10),byrow=T,nrow=3)
> b <- matrix(c(1,2,3),ncol=1)
> v <- solve(M,b)</pre>
     [,1] [,2] [,3]
> solve(M)
            Γ,17
                       \lceil,2\rceil \lceil,3\rceil
[1,] -0.6666667 -1.333333
[2,] -0.6666667 3.666667
[3,] 1.0000000 -2.000000
                                1
> V
            [,1]
[1,] -0.3333333
      0.6666667
[3,] 0.0000000
> M %*% v
     [,1]
[1,]
[2,]
[3,]
```

Case 1b: s = n and singular

• R example:

$$x + 2y + 3z = 1$$

$$4x + 5y + 6z = 2$$

$$7x + 8y + 9z = 3$$

- ▶ solve using qr gives a particular solution if it exists
- ▶ Add any multiple of kernel for general solution

```
Solution: \mathbf{v} = \mathbf{v}_p + c\mathbf{z}
\mathbf{v}_p = \begin{pmatrix} 0 \\ 0 \\ 1/3 \end{pmatrix}, \ \mathbf{z} = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}
```

```
> M1 <- matrix(1:9,byrow=T,nrow=3)</pre>
> b1 <- matrix(c(1,2,3),ncol=1)
> v1 <- solve(M1,b1)</pre>
Error in solve.default(M1, b1) :
  system is computationally singular: reciprocal condition number =
1.54198e-18
> v1_p <- qr.solve(M1,b1)</pre>
Error in qr.solve(M1, b1): singular matrix 'a' in solve
> M1
[,1] [,2] [,3]
[1,] 1 2 3
[2,] 4 5 6
[3,] 7 8 9
> det(M1)
[1] 6.661338e-16
> qr(M1)$rank
Γ17 2
> v1_p <- solve(gr(M1,LAPACK=TRUE),b1)</pre>
> v1_p
            Γ,17
Γ1.7 -0.1147976
[2,] 0.2295953
[3,] 0.2185357
> M1 %*% v1_p
     [,1]
[1,]
       1
[2,]
```

Case 2: s > n

- When there are **more equations than unknowns**, then there are **no solutions** for at least some values of **b**.
- Because a "smaller" space is going into a "bigger" one, some vectors **b** in the target space **cannot be reached** from any vector in the "smaller" space. To get technical, from the Fundamental Theorem of Linear Transformations,

$$\dim(\operatorname{Im} M) = r = n - \dim(\ker M) < n < s$$

 Although there is no solution in general, some special points (in Im M) may yield solutions.

Case 2: s > n

• R example:

$$x + 2y = -1$$
$$3x + 4y = 1$$
$$5x + 6y = 3$$
$$7x + 8y = 5$$

- → qr.solve(M,b) gives particular solution if it exists
- ▶ Warning: need to check answer since it also gives values when no solution exists(!)

Solution:

$$\mathbf{v}_p = \begin{pmatrix} 3 \\ -2 \end{pmatrix}$$

```
> M2 <- matrix(1:8,byrow=T,nrow=4)
> b2 <- matrix(c(-1,1,3,5),nrow=4)
> v2_p <- gr.solve(M2,b2)</pre>
> M2
     [,1] [,2]
[2,] 3
[3,] 5
[4,] 7
> v2_p
     [,1]
[1,]
      3
[2,] -2
> b2 <- matrix(c(-1,1,3,6),nrow=4)
> v2_not <- gr.solve(M2,b2)</pre>
> v2 not
      Γ,17
[1,] 3.50
[2,] -2.35
> M2 %*% v2_not
     [,1]
\lceil 1, \rceil -1.2
[2,] 1.1
[3,] 3.4
[4,] 5.7
```

Case 3: s < n

- When there are more unknowns than equations, then there are multiple solutions.
- Because a "bigger" space is going into a "smaller" one, some vectors must be mapped to zero. To get technical, from the Fundamental Theorem of Linear Transformations,

$$\dim(\ker M) = n - \dim(\operatorname{Im} M) > n - s > 0$$

• To any "particular solution" can be added any element of the kernel.

Case 3: s < n

• R example:

$$x + 2y + 3z + 4w = 1$$
$$5x + 6y + 7z + 8w = 1$$

- → qr.solve(M,b) gives a particular solution
- ▶ ker(M) gives (sometimes inconvenient) basis for kernel

```
Solution: \mathbf{v} = \mathbf{v}_p + c_1 \mathbf{z}_1 + c_2 \mathbf{z}_2, where \mathbf{v}_p = \begin{pmatrix} -1\\1\\0\\0 \end{pmatrix}, \ker M = \operatorname{span} \left\{ \mathbf{z}_1 = \begin{pmatrix} 2\\-3\\0\\1 \end{pmatrix}, \mathbf{z}_2 = \begin{pmatrix} 1\\-2\\1\\0 \end{pmatrix} \right\}
```

```
> M3 <- matrix(1:8,byrow=T,nrow=2)
> b3 <- matrix(c(1,1),ncol=1)
> v3_p <- gr.solve(M3,b3)</pre>
     [,1] [,2] [,3] [,4]
> v3_p
     [,1]
Γ1,
[2,]
[3,]
[4,]
> ker(M3)
           Γ,17
                       [,2]
      0.0000000 -0.5477226
      0.4082483 0.7302967
     -0.8164966 0.1825742
[4,] 0.4082483 -0.3651484
```

Case 3: s < n

• R example:

$$x + 2y + 3z + 4w = 1$$
$$5x + 6y + 7z + 8w = 1$$

- → qr.solve(M,b) gives a particular solution
- ▶ ker(M) gives (sometimes inconvenient) basis for kernel