MATA33 Lecture Notes *

Aaron Chow

Updated: 2013/08/01, 15:51:56

*These notes are mainly based on the textbook *Introductory Mathematical Analysis for Business, Economics, and the Life and Social Sciences, 13th ed.*, by Haeussler, Paul, and Wood: Pearson Education These notes are for reference only and not for sale. The author takes no responsibility for any errors; always refer back to the textbook.

Lecture Date: May 7, 2013

Linear Programming

Linear Inequalities in Two Variables

Recall the equality in the variables x and y

$$ax + by + c = 0$$

where a, b, and c are constants. This is a line in the coordinate plane.

Definition (Linear inequality)**.** A *linear inequality* in the variables x and y is an inequality that can be written in one of the forms

$$ax + by + c < 0 \qquad ax + by + c \leq 0 \qquad ax + by + c > 0 \qquad ax + by + c \geq 0$$

where a, b, and c are constants and not both a and b are zero.

Example. The following are linear inequalities in x and y:

- $2x + 3y \leq 60$,
- $5x - y > 20$,
- $x - 3y < 0$,
- $7y \geq 11$,
- $3x < 2$.

Lecture Date: May 9, 2013

Example. Suppose a consumer receives a fixed income of \$60 per week and purchases x units of product A at \$2 per unit and y units of product B at \$3 per unit. His cost will be

$$2x + 3y.$$

If he uses *all* of his \$60, then x and y satisfy

$$2x + 3y = 60 \quad x, y \geq 0;$$

this is called a *budget equation* and its graph is called a *budget line.*

If he uses *only some* of his \$60, then x and y satisfy

$$2x + 3y \leq 60 \quad x, y \geq 0.$$

Any pair (x, y) that satisfy this inequality is a solution: for example, $(15, 10)$, $(7, 11)$, etc. In general, the solution is not a line, but a *region* of the coordinate plane.

Geometrically, a (non-vertical) line $y = mx + b$ separates the coordinate plane into three parts:

1. the line itself – i.e. all points (x, y) that satisfy $y = mx + b$,
2. the region above the line (an *open half-plane*) – i.e. all points (x, y) that satisfy $y > mx + b$, and
3. the region below the line (also an *open half-plane*) – i.e. all points (x, y) that satisfy $y < mx + b$.

Therefore, the solution to:

- $y > mx + b$ is the open half-plane above the line,
- $y < mx + b$ is the open half-plane below the line,
- $y \geq mx + b$ is the *closed half-plane* consisting of the open half-plane above the line together with the line itself, and
- $y \leq mx + b$ is the *closed half-plane* consisting of the open half-plane below the line together with the line itself.

A vertical line $x = a$ separates the coordinate plane into three parts:

1. the line itself – i.e. all points (x, y) that satisfy $x = a$,
2. the open half-plane to the right the line – i.e. all points (x, y) that satisfy $x > a$, and
3. the open half-plane to the left the line – i.e. all points (x, y) that satisfy $x < a$.

Example. Solve the inequality

$$2(2x - y) < 2(x + y) - 4.$$

Solution: First, "isolate for y":

$$\begin{aligned} 2(2x - y) &< 2(x + y) - 4 \\ 4x - 2y &< 2x + 2y - 4 \\ 2x - 4y + 4 &< 0 \\ y &> \frac{1}{2}x + 1. \end{aligned}$$

Therefore, the solution is the open half-plane above the line $y = \frac{1}{2}x + 1$.

Systems of Linear Inequalities in Two Variables

The solution to a *system of linear inequalities*, e.g.

$$\begin{cases} a_1x + b_1y + c_1 < 0 \\ a_2x + b_2y + c_2 \leq 0 \\ a_3x + b_3y + c_3 > 0 \\ a_4x + b_4y + c_4 \geq 0, \end{cases}$$

consists of all points (x, y) that satisfy *all* inequalities *simultaneously.*

Geometrically, the solution (a region) to a system of linear inequalities is the intersection of the solutions of each linear inequality, that is, the region common to all the regions determine by each individual inequality.

Example. Solve the system of inequalities

$$\begin{cases} 2x + y \geq 10 \\ x - y - 2 \leq 0 \end{cases}$$

Solution: First, we solve each given linear inequality as before:

$$\begin{cases} y \geq -2x + 10 \\ y \geq x - 2 \end{cases}$$

Therefore, the solution is the intersection of:

1. the closed half-plane above the line $y = -2x + 10$, and
2. the closed half-plane above the line $y = x - 2$.

Exercise. Solve the system of inequalities

$$\begin{cases} 2x + y \geq 10 \\ x - y - 2 \leq 0 \\ x \geq 0 \\ y \geq 0. \end{cases}$$

Solution: First, we solve each given linear inequality as before:

$$\begin{cases} y \geq -2x + 10 \\ y \geq x - 2 \\ x \geq 0 \\ y \geq 0 \end{cases}$$

Therefore, the solution is the intersection of:

1. the region of the previous example,
2. the closed half-plane to the right of the line $x = 0$, and
3. the closed half-plane above the line $y = 0$.

Example. A company produces two products, A and B. Product A takes 3 worker-hours to assemble and 0.5 worker-hour to paint. Product B takes 2 worker-hours to assemble and 1 worker-hour to paint. The daily maximum number of worker-hours available for assembly is 240, and the daily maximum number of worker-hours available for painting is 80.

Write a system of linear inequalities to describe the situation and find the region described by this system.

Solution: Let x and y denote the number of products A and B produced per day respectively. Then x and y satisfy the system

$$\begin{cases} 3x + 2y \leq 240 \\ \frac{1}{2}x + y \leq 80 \\ x \geq 0 \\ y \geq 0. \end{cases}$$

The region is the intersection of the regions described by each inequality.

Lecture Date: May 14, 2013

Linear Programming

Linear programming is about maximizing or minimizing and *objective function* subject to *constraints.*

Definition (Linear Function)**.** A *linear function in x and y* is a function of two variables of the form

$$P(x, y) = ax + by$$

where a and b are constants.

Remark. The domain of a linear function as described above is $\mathbb{R}^{\nvDash} = \{(x, y) : x, y \in \mathbb{R}\} = (-\infty, \infty) \times (-\infty, \infty)$. However, in applications, x and y will often be positive or even integers (e.g. quantities). Thus, the domain is immediately restricted to $\{(x, y) : x, y \in \mathbb{R}, x, y \geq 0\} = [0, \infty) \times [0, \infty)$, or perhaps further restricted if there are additional constraints.

Definition (Optimization)**.** To *optimize* a function is to either maximize or minimize it depending on the nature of the problem.

Definition (Objective Function)**.** The function to be optimized is called teh *objective function.*

Definition (Feasible Region, Feasible Points)**.** The region defined by the system of linear constraints is called the *feasible region.*

A point in the feasible region is called a *feasible point.*

Remark. By definition, the set of all feasible points is the feasible region.

Example. Suppose a company produces two products, A and B, in three stages using three machines. The amount of time needed on each machine, the hours available on each machine, and the profit for selling A and B are summarized in the following table:

	Product A	Product B	Hours Available
Machine 1	2 hr	1 hr	180 hr
Machine 2	1 hr	2 hr	160 hr
Machine 3	1 hr	1 hr	100 hr
Profit	\$4	\$6	

How many of each product should the company produce to maximize profit?

Solution: Let x and y denote the number of products A and B the company produces respectively.

Then x and y must satisfy

$$\begin{cases} x \geq 0 \\ y \geq 0 \\ 2x + y \leq 180 \\ x + 2y \leq 160 \\ x + y \leq 100. \end{cases}$$

The *profit function* (i.e. the objective function) is

$$P = P(x, y) = 4x + 6y.$$

Thus, we want to *maximize* the objective function

$$P = 4x + 6y.$$

subject to the system of constraints

$$\begin{cases} x \geq 0 \\ y \geq 0 \\ 2x + y \leq 180 \\ x + 2y \leq 160 \\ x + y \leq 100. \end{cases}$$

Now, each specific value of P defines a particular line, an *isoprofit line*. Indeed, the profit function can be written as

$$y = -\frac{3}{2}x + \frac{P}{6}.$$

Mathematically, this is a *family* of curves, and each specialization of P gives a *level curve*. Note the slope is independent of P, and that a larger P corresponds to a higher line.

Graphing the feasible region given by the constraints, we see the highest line passes through the point $(40, 60)$ where

$$P = 4(40) + 6(60) = 520.$$

Therefore the company should produce 40 units of product A and 60 units of product B for a maximum profit of \$520.

In general, we will *optimize* a linear *objective function* subject to a system *linear constraints*. To solve the problem geometrically, we find the *feasible region* determined by the constraints and look for the highest (or lowest) possible level curve.

Definition (Boundedness)**.** If a feasible region is contained in a circle, then it is a *bounded feasible region*; otherwise, it is *unbounded*.

Definition (Emptiness)**.** If a feasible region contains at least a point, then it is said to be *nonempty*; otherwise, it is said to be *empty*.

Remark. The above definitions are also for a general two-dimensional subset.

Theorem (Extreme Value Theorem)**.** *A linear function defined on a nonempty bounded feasible region has both a maximum and minimum value, and this value is attained at a corner point.*

Remark. A *corner point* occurs when two constraint boundaries meet. Therefore, to find the coordinates of a corner point, we just solve a 2×2 linear system.

Exercise. Google “extreme value theorem”.

Lecture Date: May 16, 2013

Example. Consider the previous example. We can apply the theorem to find the optimum point; just evaluate P at each corner point.

One corner point occurs at the intersection of

$$\begin{cases} x + 2y = 160 \\ x + y = 100. \end{cases}$$

The solution to this system of equalities is $(40, 60)$.

The five corner points and the values of P are:

$$P(40,60) = 520 \quad P(80,20) = 440 \quad P(90,0) = 360 \quad P(0,0) = 0 \quad P(0,80) = 480.$$

We see that the maximum occurs at $(40, 60)$ and the maximum value of P is 520.

Exercise. Find the maximum and minimum of

$$P = 3x + y$$

subject to

$$\begin{cases} 2x + y \leq 8 \\ 2x + 3y \leq 12 \\ x, y \geq 0. \end{cases}$$

Solution: The feasible region is nonempty and bounded, so by the theorem, the maximum is attained at a corner point.

The four corner points and the values of P are:

$$P(0,0) = 0 \qquad P(4,0) = 12 \qquad P(3,2) = 11 \qquad P(0,4) = 4.$$

Therefore the maximum value is 12 attained at $(4, 0)$, and the minimum value is 0 attained at $(0, 0)$.

Empty Feasible Region

Example (Empty Feasible Region)**.** Minimize

$$Z = 8x - 3y$$

subject to

$$\begin{cases} -x + 3y = 21 \\ x + y \leq 5 \\ x, y \geq 0. \end{cases}$$

Solution: Note that the feasible region is empty since the line (equality) $-x + 3y = 21$ lies outside the region defined by the rest of the constraint system.

Therefore the problem has no solution.

Remark. In general, whenever the feasible region of a linear programming problem is empty, the problem has no optimum solution.

Unbounded Feasible Region

Example. Optimize

$$Z = x + y$$

subject to

$$\begin{cases} y = 2 \\ x, y \geq 0. \end{cases}$$

Solution: : In this case, the feasible region is just the infinite horizontal line $y = 2$ in the first quadrant. The objective function can be rewritten as

$$Z = x + 2 \quad x \geq 0;$$

this satisfy all the constraints.

It is clear that Z can get arbitrarily large as x increases without bounds. Therefore there is no maximum.

However, there is a minimum; it occurs at $(0, 2)$ and the minimum value of Z is 2.

Example. A manufacture needs to purchase parts A, B, and C for its production. It requires at least 160 units of A, 200 units of B, and 80 units of C. The manufacture can purchase these parts in boxes from companies X and Y. A box from company X, which costs \$8 per box, contains 3 units of A, 5 units of B, and 1 unit of C; a box from company Y, which costs \$6 per box, contains 2 units of A, 2 units of B, and 2 units of C. How many boxes should the manufacture buy from each company to minimize cost?

Solution: We summarize the information as follows:

	Company X's Box	Company Y's Box	Minimum Units Required
Part A	3 units	2 units	160 units
Part B	5 units	2 units	200 units
Part C	1 unit	2 units	80 units
Cost	\$8	\$6	

Let x and y denote the number of boxes purchased from companies X and Y respectively. Since at least 160 units of part A is required, x and y must satisfy

$$3x + 2y \geq 160.$$

There are similar inequalities are determined by requirements on parts B and C. Thus, feasible region is defined by the system of linear inequalities:

$$\begin{cases} 3x + 2y \geq 160 \\ 5x + 2y \geq 200 \\ x + 2y \geq 80 \\ x, y \geq 0. \end{cases}$$

This region is *unbounded*.

We wish to *minimize* the cost function

$$C = C(x, y) = 8x + 6y$$

subject to the constraints above.

The minimum occurs at the corner point $(40, 20)$ where $C = 440$.

Therefore the manufacturer should purchase 40 boxes from company X and 20 boxes from company Y.

Example. The linear programming problem in the previous example *has no maximum*!

If we had followed the algorithm blindly, we would evaluate the cost function at the corner points:

$$C(80, 0) = 640 \qquad C(40, 20) = 440 \qquad C(20, 50) = 460 \qquad C(0, 100) = 600,$$

and conclude that the maximum value of C is 640 occurring at $(80, 0)$.

This is, of course, absurd because the manufacturer can just by many many boxes without regard to cost until the minimum quantities are met!

Multiple Optimal Solutions

Definition (Multiple Optimal Solutions)**.** If the objection function in a linear programming problem attains its optimum value at more than one feasible point, then *multiple optimal solutions* is said to exist.

Example. Maximize

$$Z = 2x + 4y$$

subject to

$$\begin{cases} x - 4y \leq -8 \\ x + 2y \leq 16 \\ x, y \geq 0 \end{cases}$$

Solution: The three corner points and the values of Z are:

$$Z(0, 2) = 8 \qquad Z(8, 4) = 32 \qquad Z(0, 8) = 32.$$

We see that Z attains its maximum at both $(8, 4)$ and $(0, 8)$. In fact, Z attains its maximum at all the points on the line segment connecting $(8, 4)$ and $(0, 8)$.

Theorem. *If (x_1, y_1) and (x_2, y_2) are two corner points at which an objective function is optimum, then the function will also be optimum at all points on the line segment connecting (x_1, y_1) and (x_2, y_2), i.e. the set*

$$\{(x, y) \in \mathbb{R}^2 : x = (1 - t)x_1 + tx_2, y = (1 - t)y_1 + ty_2, 0 \leq t \leq 1\}$$

Example. We check the theorem is true with the previous example.

Here, the corner points are

$$(x_1, y_1) = (8, 4) \qquad (x_2, y_2) = (0, 8)$$

and any point (x, y) on the line segment joining them has

$$x = (1 - t)x_1 + tx_2 = (1 - t)(8) + t(0) = 8(1 - t)$$
$$y = (1 - t)y_1 + ty_2 = (1 - t)(4) + t(8) = 4(1 + t)$$

with $0 \leq t \leq 1$.

Indeed, substituting $(x, y) = (8(1 - t), 4(1 + t))$ into $Z = 2x + 4y$ gives

$$Z = 2(8(1 - t)) + 4(4(1 + t)) = 16 - 16t + 16 + 16t = 32.$$

Matrix Algebra

"Matrices are arrays of numbers."

Matrices

Definition (Matrices)**.** An $m \times n$ *matrix* is a rectangular array A of numbers with m rows and n columns,

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

The *size of* A is $m \times n$. The numbers a_{ij} are the *entries* of A, where the subscripts i and j denote the row and column subscripts.

Remark. Sometimes, we write $A = [a_{ij}]$ to denote a matrix with entries a_{ij}.

Example.

1. $\begin{bmatrix} 3 & 1 & 0 \\ -5 & 2 & 5 \\ 34 & -7 & 1 \end{bmatrix}$ is a 3×3 matrix.
2. $\begin{bmatrix} 3 & 1 & 0 \\ -5 & 2 & 5 \end{bmatrix}$ is a 2×3 matrix.
3. $\begin{bmatrix} 3 & 1 & 0 \end{bmatrix}$ is a 1×3 matrix.
4. $\begin{bmatrix} 3 \\ -5 \\ 34 \end{bmatrix}$ is a 3×1 matrix.
5. $\begin{bmatrix} 3 \end{bmatrix}$ is a 1×1 matrix.

Definition (Row and Column Vector)**.** A *row vector* is a matrix with one row:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \end{bmatrix}.$$

Its size is $1 \times n$.

A *column vector* is a matrix with one column:

$$\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}.$$

Its size is $m \times 1$.

Example (Working with entries)**.** If $A = [a_{ij}]$ is 3×4 and $a_{ij} = i + j$, find A.

Solution: Since A is 3×4, we have

$$i = 1, 2, 3 \qquad\qquad j = 1, 2, 3, 4.$$

Also, $a_{ij} = i + j$, so, for example, $a_{11} = 1 + 1 = 2$, $a_{21} = 2 + 1 = 3$, etc. Therefore

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} = \begin{bmatrix} 1+1 & 1+2 & 1+3 & 1+4 \\ 2+1 & 2+2 & 2+3 & 2+4 \\ 3+1 & 3+2 & 3+3 & 3+4 \end{bmatrix} = \begin{bmatrix} 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 7 \end{bmatrix}$$

Example (Storing data in matrices)**.** Suppose a company sells 101 units of product A at \$1/unit, 202 units of product B at \$2/unit, and 303 units of product C at \$3/unit.

We can represent the data matrices. For example, the number of units sold can be represented by the row vector $\begin{bmatrix} 101 & 202 & 303 \end{bmatrix}$ and the prices per unit by the column vector $\begin{bmatrix} \$1 \\ \$2 \\ \$3 \end{bmatrix}$.

Lecture Date: May 21, 2013

Equality of Matrices

Definition (Equality of Matrices)**.** Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be matrices. A and B are *equal* if they have the same size and the entries $a_{ij} = b_{ij}$ for all i and j. In this case, we write $A = B$.

Example.

$$\begin{bmatrix} 1+1 & 1+2 & 1+3 & 1+4 \\ 2+1 & 2+2 & 2+3 & 2+4 \\ 3+1 & 3+2 & 3+3 & 3+4 \end{bmatrix} = \begin{bmatrix} 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 7 \end{bmatrix}$$

but

$$\begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \neq \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \qquad \begin{bmatrix} 1 & 1 \end{bmatrix} \neq \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \neq \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \qquad \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \neq \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}.$$

Example. Solve for x, y, and z if

$$\begin{bmatrix} x & y \\ x+y & z \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 5 & 7 \end{bmatrix}.$$

Solution: Equating the entries, we have

$$\begin{cases} x = 2 \\ y = 3 \\ x + y = 5 \\ z7. \end{cases}$$

Thus, $x = 2$, $y = 3$, and $z = 7$.

Transpose of a Matrix

Definition (Transpose of Matrices)**.** If $A = [a_{ij}]$, then the *transpose* of A is $A^T = [a_{ji}]$.
In other words, A^T is obtained from A by interchanging its rows with its columns.

Remark. If A is $m \times n$, then A^T is $n \times m$.

Example.

1. If $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ then $A^T = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$.
2. $\begin{bmatrix} 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 7 \end{bmatrix}^T = \begin{bmatrix} 2 & 3 & 4 \\ 3 & 4 & 5 \\ 4 & 5 & 6 \\ 5 & 6 & 7 \end{bmatrix}.$

3. $\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix}^T = \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \\ a_{143} & a_{24} & a_{34} \end{bmatrix}.$

4. $\begin{bmatrix} 2 & 3 & 5 & 7 \end{bmatrix}^T = \begin{bmatrix} 2 \\ 3 \\ 5 \\ 7 \end{bmatrix}.$

Remark. $(A^T)^T = A$.

Special Matrices

Definition (Zero Matrix)**.** An $m \times n$ matrix whose entries are all zero is called the $m \times n$ *zero matrix*. It is written $0_{m\times n}$ or just 0 if the size is clear from context.

Example.

1. $0_{2\times 2} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ is the 2×2 zero matrix.

2. $0_{3\times 4} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ is the 3×4 zero matrix.

Definition (Square Matrix)**.** A *square matrix* is a matrix that has the same number of columns as rows.

Definition (Order)**.** If a square matrix A has size $n \times n$, then we say *A is of order n*.

Definition (Main Diagonal)**.** Let $A = [a_{ij}]$ be a square matrix of order n. Then the entries $a_{11}, a_{22}, \ldots, a_{nn}$ lie on the *main diagonal*.

Example. Consider the matrix

$$A = \begin{bmatrix} \mathbf{1} & 2 & 3 \\ 4 & \mathbf{5} & 6 \\ 7 & 8 & \mathbf{9} \end{bmatrix}.$$

A is a square matrix of order 3. The bolded entries lie on the main diagonal.

Definition (Diagonal Matrix)**.** A square matrix $A = [a_{ij}]$ is called a *diagonal matrix* if all the entries off the main diagonal are zero, i.e. $a_{ij} = 0$ if $i \neq j$.

Example. The following are diagonal matrices:

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 8 \end{bmatrix} \quad B = \begin{bmatrix} -8 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{bmatrix} \quad C = \begin{bmatrix} 87 & 0 & 0 & 0 \\ 0 & 2013 & 0 & 0 \\ 0 & 0 & -734 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Definition (Upper and Lower Triangular Matrix)**.** A square matrix $A = [a_{ij}]$ is called an *upper triangular matrix* if all the entries below the main diagonal are zero, i.e. $a_{ij} = 0$ if $i > j$.

A square matrix $B = [b_{ij}]$ is called a *lower triangular matrix* if all the entries above the main diagonal are zero, i.e. $a_{ij} = 0$ if $i < j$.

A square matrix is *triangular* if it is either upper or lower triangular.

Remark. Diagonal matrices are triangular; in fact, they are *both* upper *and* lower triangular.

Example. The following are upper triangular matrices:

$$A = \begin{bmatrix} 1 & -2 \\ 0 & 8 \end{bmatrix} \quad B = \begin{bmatrix} -8 & 11 & -2 \\ 0 & 3 & 1 \\ 0 & 0 & 4 \end{bmatrix} \quad C = \begin{bmatrix} 87 & -55 & 31 & 99 \\ 0 & 2013 & -88 & 7 \\ 0 & 0 & -734 & 2 \\ 0 & 0 & 0 & 56 \end{bmatrix} \quad D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Example. The following are lower triangular matrices:

$$E = \begin{bmatrix} 1 & 0 \\ 5 & 8 \end{bmatrix} \quad F = \begin{bmatrix} -8 & 0 & 0 \\ 5 & 3 & 0 \\ 8 & 3 & 4 \end{bmatrix} \quad G = \begin{bmatrix} 87 & 0 & 0 & 0 \\ -55 & 2013 & 0 & 0 \\ 35 & 42 & -734 & 0 \\ 87 & 0 & 1 & 0 \end{bmatrix} \quad H = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Remark. The transpose of an upper triangular matrix is lower triangular, and conversely, the transpose of a lower triangular matrix is upper triangular.

Remark. Diagonal matrices are both upper and lower triangular matrices.

Lecture Date: May 23, 2013

Matrix Addition and Scalar Multiplication

Addition

Definition (Matrix Sum)**.** Given two $m \times n$ matrices A and B, the *sum* $A + B$ is the $m \times n$ matrix obtained by adding the entries of A and B. That is, if $A = [a_{ij}]$ and $B = [b_{ij}]$ then $A + B = [a_{ij} + b_{ij}]$.

If the sizes of A and B are different, then $A + B$ is not defined.

Example.

$$\begin{bmatrix}1 & 0\\3 & 2\end{bmatrix} + \begin{bmatrix}2 & -2\\7 & 0\end{bmatrix} = \begin{bmatrix}3 & -2\\10 & 2\end{bmatrix} \qquad \begin{bmatrix}1 & 2\\3 & 4\\5 & 6\end{bmatrix} + \begin{bmatrix}7 & -1\\5 & 0\\2 & -3\end{bmatrix} = \begin{bmatrix}8 & 1\\8 & 4\\7 & 3\end{bmatrix}$$

Example. The following sums are not defined since the sizes of the matrices are different:

$$\begin{bmatrix}1 & 0\\3 & 2\end{bmatrix} + \begin{bmatrix}2\\7\end{bmatrix} \qquad \begin{bmatrix}1 & 2\\3 & 4\\5 & 6\end{bmatrix} + \begin{bmatrix}7 & -1 & 3\\5 & 0 & -9\end{bmatrix} \qquad \begin{bmatrix}1 & 2 & 3\end{bmatrix} + \begin{bmatrix}2\\3\\7\end{bmatrix}.$$

Theorem (Properties of Matrix Addition)**.** *Suppose A, B, C, and 0 (zero matrix) are matrices with the same size. Then*

- $A + B = B + A$ *(commutative property),*
- $A + (B + C) = (A + B) + C$ *(associative property), and*
- $A + 0 = A = 0 + A$ *(identity property).*

Exercise. Let

$$A = \begin{bmatrix}1 & 2 & 3\\4 & 5 & 6\end{bmatrix} \qquad B = \begin{bmatrix}3 & 0 & 5\\-4 & 1 & 2\end{bmatrix} \qquad C = \begin{bmatrix}7 & -1 & -4\\-1 & 2 & 5\end{bmatrix} \qquad 0 = \begin{bmatrix}0 & 0 & 0\\0 & 0 & 0\end{bmatrix}.$$

Check that the three properties of the theorem, i.e.

$$A + B = B + A \qquad A + (B + C) = (A + B) + C \qquad A + 0 = A = 0 + A$$

Example. Suppose the world consists of four products: A, B, C, and D. Demand for them come from four consumers (C_1, C_2, C_3, and C_4) and four manufactures (M_1, M_2, M_3, and M_4). The demand from the consumers can be described by the following vectors:

$$D_{C_1} = \begin{bmatrix}2 & 4 & 1 & 3\end{bmatrix} \qquad D_{C_2} = \begin{bmatrix}3 & 0 & 5 & 2\end{bmatrix}$$
$$D_{C_3} = \begin{bmatrix}1 & 8 & 3 & 5\end{bmatrix} \qquad D_{C_4} = \begin{bmatrix}6 & 2 & 0 & 4\end{bmatrix}.$$

Likewise, the demand from the manufactures can be described by the following vectors:

$$D_{M_1} = \begin{bmatrix}10 & 12 & 9 & 7\end{bmatrix} \qquad D_{M_2} = \begin{bmatrix}6 & 8 & 11 & 8\end{bmatrix}$$
$$D_{M_3} = \begin{bmatrix}16 & 7 & 8 & 11\end{bmatrix} \qquad D_{M_4} = \begin{bmatrix}3 & 7 & 9 & 13\end{bmatrix}.$$

The total consumer demand is therefore

$$D_C = D_{C_1} + D_{C_2} + D_{C_3} + D_{C_4} = \begin{bmatrix} 12 & 14 & 9 & 14 \end{bmatrix}$$

and the total manufacturing demand is

$$D_M = D_{M_1} + D_{M_2} + D_{M_3} + D_{M_4} = \begin{bmatrix} 35 & 34 & 37 & 39 \end{bmatrix}.$$

The overall demand is then

$$D_C + D_M = \begin{bmatrix} 47 & 48 & 46 & 53 \end{bmatrix}.$$

Scalar Multiplication

Definition. Matrix Scalar Multiplication If A is an $m \times n$ matrix and $k \in \mathbb{R}$, then the *scalar multiple* kA is the $m \times n$ matrix obtained by multiplying the entries of A by k. That is, if $A = [a_{ij}]$ then $kA = [ka_{ij}]$.

Example.

$$4\begin{bmatrix} 1 & 0 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 12 & 8 \end{bmatrix} \qquad -3\begin{bmatrix} 1 & -2 \\ 3 & 4 \\ -5 & 6 \end{bmatrix} = \begin{bmatrix} -3 & 6 \\ -9 & -12 \\ 15 & -18 \end{bmatrix} \qquad 0\begin{bmatrix} 3 & -1 \\ 23 & -17 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Remark. $0A = 0$ for *any* matrix A.

Exercise. Let

$$A = \begin{bmatrix} 1 & 2 \\ 5 & 4 \end{bmatrix} \qquad B = \begin{bmatrix} -5 & 0 \\ -2 & 3 \end{bmatrix} \qquad C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Compute

1. $5A$,
2. $-\frac{3}{7}B$,
3. $\frac{1}{3}C - 2B + \frac{2}{5}A$, and
4. kC for any $k \in \mathbb{R}$.

Theorem (Properties of Scalar Multiplication)**.** *Suppose A and B are matrices with the same size, and $c, d \in \mathbb{R}$ are scalars. Then*

- $c(A + B) = cA + cB$,
- $(c + d)A = cA + dA$,
- $c(dA) = (cd)A$,
- $0A = 0$, *and*
- $c0 = 0$.

Exercise. Let

$$A = \begin{bmatrix} 1 & 2 \\ 5 & 4 \end{bmatrix} \qquad B = \begin{bmatrix} -5 & 0 \\ -2 & 3 \end{bmatrix}.$$

Verify that

1. $3(A+B) = 3A + 3B$,
2. $9A = (2+7)A = 2A + 7A$, and
3. $2(5A) = 10A$.

Theorem (Transposition)**.** *Suppose A and B are matrices with the same size, and $k \in \mathbb{R}$ is a scalar. Then*

$$(A+B)^T = A^T + B^T \qquad (kA)^T = kA^T.$$

Subtraction

Definition (Matrix Negation)**.** If A is a matrix, then the scalar multiple $(-1)A$, written just $-A$, is the *negative of* A.

Example. If $A = [a_{ij}]$, then $-A = [-a_{ij}]$

Definition (Matrix Subtraction)**.** Given two $m \times n$ matrices A and B, then we define

$$A - B = A + (-B).$$

Example.

$$\begin{bmatrix} 1 & 0 \\ 3 & 2 \end{bmatrix} - \begin{bmatrix} 2 & -2 \\ 7 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 2 \\ -4 & 2 \end{bmatrix} \qquad \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} - \begin{bmatrix} 7 & -1 \\ 5 & 0 \\ 2 & -3 \end{bmatrix} = \begin{bmatrix} -6 & 3 \\ -2 & 4 \\ 3 & 9 \end{bmatrix}$$

Example. Suppose a manufacturer produces and sells three products, A, B, and C, at \$10, \$3, and \$7 per unit respectively. The manufacturing costs are \$5, \$1, and \$4 respectively. Represent the price, cost, and profit per unit as column vectors.

Solution: The price/unit can be represented as the vector $\begin{bmatrix} \$10 \\ \$3 \\ \$7 \end{bmatrix}$. The cost/unit can be represented as the vector $\begin{bmatrix} \$5 \\ \$1 \\ \$4 \end{bmatrix}$. The profit/unit can be represented as the vector difference $\begin{bmatrix} \$10 \\ \$3 \\ \$7 \end{bmatrix} - \begin{bmatrix} \$5 \\ \$1 \\ \$4 \end{bmatrix} = \begin{bmatrix} \$5 \\ \$2 \\ \$3 \end{bmatrix}$.

Matrix Multiplication

So far, we know what $A + B$ and kB means. We know define yet another operation, *matrix multiplication*, not to be confused with *scalar multiplication*.

Definition (Matrix Multiplication)**.** Let A be an $m \times n$ matrix and B be an $n \times p$ matrix. Then the *matrix product* AB is an $m \times p$ matrix whose entires are obtained by multiplying each entry of the rows of A by the each entry of the columns of B. That is, if $A = [a_{ij}]$ and $B = [b_{jk}]$, then the ik-th entry of $AB = C = [c_{ik}]$ is given by

$$c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk} = a_{i1} b_{1k} + a_{i2} b_{2k} + \cdots + a_{in} b_{nk}.$$

If the number of columns of A does not equal the number of rows of B, then AB is not defined.

Example. If $A = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$ and $B = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$, then AB is a 1×1 matrix and

$$AB = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} (1)(4) + (2)(5) + (3)(6) \end{bmatrix} = \begin{bmatrix} 32 \end{bmatrix}.$$

On the other hand, BA is a 3×3 matrix and

$$BA = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} = \begin{bmatrix} (4)(1) & (4)(2) & (4)(3) \\ (5)(1) & (5)(2) & (5)(3) \\ (6)(1) & (6)(2) & (6)(3) \end{bmatrix} = \begin{bmatrix} 4 & 8 & 12 \\ 5 & 10 & 15 \\ 6 & 12 & 18 \end{bmatrix}.$$

Example. Let

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \qquad B = \begin{bmatrix} -2 & 1 \\ 0 & 3 \end{bmatrix}.$$

Compute and compare AB and BA.

Solution:

$$AB = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} -2 & 1 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} (1)(-2) + (2)(0) & (1)(1) + (2)(3) \\ (3)(-2) + (4)(0) & (3)(1) + (4)(3) \end{bmatrix} = \begin{bmatrix} -2 & 7 \\ -6 & 15 \end{bmatrix}$$

and

$$BA = \begin{bmatrix} -2 & 1 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} (-2)(1) + (1)(3) & (-2)(2) + (1)(4) \\ (0)(1) + (3)(3) & (0)(2) + (3)(4) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 9 & 12 \end{bmatrix}.$$

AB and BA are both 2×2 matrices, but $AB \neq BA$.

Remark. In general, matrix multiplication is *not commutative*; that is, $AB \neq BA$ (assuming both make sense to begin with).

Example.

$$\underbrace{\begin{bmatrix} -2 & 1 & 0 \\ 1 & 0 & 3 \end{bmatrix}}_{2\times 3} \underbrace{\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}}_{3\times 2} = \underbrace{\begin{bmatrix} 1 & 0 \\ 16 & 20 \end{bmatrix}}_{2\times 2} \qquad \underbrace{\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}}_{3\times 2} \underbrace{\begin{bmatrix} -2 & 1 & 0 \\ 1 & 0 & 3 \end{bmatrix}}_{2\times 3} = \underbrace{\begin{bmatrix} 0 & 1 & 6 \\ -2 & 3 & 12 \\ -4 & 5 & 18 \end{bmatrix}}_{3\times 3}.$$

Example. Recall a previous example where a company sells 101 units of product A at \$1/unit, 202 units of product B at \$2/unit, and 303 units of product C at \$3/unit.

The data were represented as matrices: the number of units sold by the row vector $Q = \begin{bmatrix} 101 & 202 & 303 \end{bmatrix}$ and the prices per unit by the column vector $P = \begin{bmatrix} \$1 \\ \$2 \\ \$3 \end{bmatrix}$.

The total revenue is given by the matrix product

$$QP = \begin{bmatrix} 101 & 202 & 303 \end{bmatrix} \begin{bmatrix} \$1 \\ \$2 \\ \$3 \end{bmatrix} = \begin{bmatrix} \$101 + \$404 + \$909 \end{bmatrix} = \begin{bmatrix} \$1414 \end{bmatrix}.$$

Think of this as the multivariate version of

$$\text{revenue} = \text{number of units sold} \times \text{price per unit}.$$

Theorem (Properties of Matrix Multiplication)**.** *Let A, B, C be matrices of the appropriate sizes, and let $k \in \mathbb{R}$ be a scalar.*

1. $A(BC) = (AB)C$ *(associative property).*
2. $A(B + C) = AB + AC$ *and* $(A + B)C = AC + BC$ *(distributive properties).*
3. $kAB = k(AB) = (kA)B = A(kB)$.
4. $(AB)^T = B^T A^T$.

Exercise. Let

$$A = \begin{bmatrix} 1 & -2 \\ -3 & 4 \end{bmatrix} \qquad B = \begin{bmatrix} 3 & 0 & -1 \\ 1 & 1 & 2 \end{bmatrix} \qquad C = \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}.$$

1. What is the size of ABC?
2. Verify that $ABC = (AB)C = A(BC)$.

Exercise. Let

$$A = \begin{bmatrix} 1 & -2 \\ -3 & 4 \end{bmatrix} \qquad B = \begin{bmatrix} 3 & 0 \\ 1 & 1 \end{bmatrix} \qquad C = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}.$$

1. Verify that $A(B + C) = AB + AC$.
2. Verify that $(A + B)C = AC + BC$.

Exercise. Let

$$A = \begin{bmatrix} 1 & -2 \\ -3 & 4 \end{bmatrix} \qquad B = \begin{bmatrix} 3 & 0 & -1 \\ 1 & 1 & 2 \end{bmatrix}$$

1. Verify that $k(AB) = (kA)(B) = A(kB)$ for any $k \in \mathbb{R}$.
2. Verify that $(AB)^T = B^T A^T$.

Definition (Identity Matrix)**.** The $n \times n$ *identity matrix*, denoted I_n (or just I if clear from context), is the diagonal matrix whose main diagonal entries are all 1's. That is,

$$I_n = \underbrace{\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}}_{n \times n}$$

Example.

$$I_1 = \begin{bmatrix} 1 \end{bmatrix} \qquad I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Example.

$$I_n^T = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}^T = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = I_n$$

Theorem. *Let A be any matrix and I the identity matrix of the appropriate size. Then*

1. *$AI = A$ and*
2. *$IA = A$.*

In particular, if A is a square matrix, then $AI = A = IA$.

Exercise. Let

$$A = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \qquad I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad 0 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Compute:

1. $I - A$,
2. $3(A - 2I)$,

3. $A0$ and $0A$, and
4. $AB(A0 + 3(A - 2I))$.

Definition (Matrix Powers)**.** If A is an $n \times n$ square matrix and p is a positive integer, then *the* $p - th$ *power of* A, written A^p, is defined as

$$A^p = \underbrace{AA \cdots A}_{p \text{ times}}.$$

If $p = 0$, define $A^0 = I$.

Example. Let $A = \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}$. Compute A^3.

Solution:

$$\begin{aligned} A^2 &= AA = \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 3 & 4 \end{bmatrix} \\ A^3 &= A^2 A = \begin{bmatrix} 1 & 0 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 7 & 8 \end{bmatrix} \end{aligned}$$

Matrix Equations

Definition (Matrix Equation)**.** An equation of the form

$$AX = B$$

is a *matrix equation*. Here A is called the *coefficient matrix*, and X and B are column matrices.

Remark. If A is $m \times n$, then X is $n \times 1$ and B is $m \times 1$.

Remark. Here, X plays the role of the variable.

Remark. $AX = B$ is the multivariate version of the linear equation $ax = b$.

Example. Consider the matrix equation

$$AX = B$$

$$\underbrace{\begin{bmatrix} 1 & 4 & -2 \\ 2 & -3 & 1 \end{bmatrix}}_{2\times 3} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}}_{3\times 1} = \underbrace{\begin{bmatrix} 4 \\ -3 \end{bmatrix}}_{2\times 1}.$$

Carrying out the multiplication on the left hand side yields

$$\begin{bmatrix} x_1 + 4x_2 - 2x_3 \\ 2x_1 - 3x_2 + x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ -3 \end{bmatrix}$$

which is equivalent to the system of equations

$$\begin{cases} x_1 + 4x_2 - 2x_3 = 4 \\ 2x_1 - 3x_2 + x_3 = -3. \end{cases}$$

Example. The system of equations

$$\begin{cases} x_1 + 2x_2 = 2 \\ -4x_1 + 6x_3 = -7 \end{cases}$$

in matrix form is

$$\begin{bmatrix} 1 & 2 \\ -4 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ -7 \end{bmatrix}.$$

The system of equations

$$\begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 = 0 \\ -x_2 + 6x_3 = 0 \\ 3x_1 - 11x_3 - 2x_4 = 0 \end{cases}$$

in matrix form is

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & 6 & 0 \\ 3 & 0 & -11 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Lecture Date: May 28, 2013

Example (General Example)**.** The system

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + a_{n2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

corresponds to the matrix equation

$$AX = B$$

where $A = [a_{ij}]$ is the coefficient matrix, $X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, and $B = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$.

Definition (Solution)**.** A *solution* to the matrix equation $AX = B$ is a column vector C (of the same size as X) such that $AC = B$.

Example. Consider again the matrix equation

$$AX = B$$

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

The vector $C = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ is a solution to $AX = B$ since

$$AC = \begin{bmatrix} 1 & 4 & -2 \\ 2 & -3 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 4 \\ -3 \end{bmatrix} = B.$$

The vector $D = \begin{bmatrix} 2 \\ 6 \\ 11 \end{bmatrix}$ is also a solution to $AX = B$ since

$$AD = \begin{bmatrix} 1 & 4 & -2 \\ 2 & -3 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 6 \\ 11 \end{bmatrix} = \begin{bmatrix} 4 \\ -3 \end{bmatrix} = B.$$

In fact, all solutions to this system have the form $\begin{bmatrix} \frac{2}{11}s \\ 1 + \frac{5}{11}s \\ s \end{bmatrix}$, $s \in \mathbb{R}$. Indeed, if we solve the corresponding system by reducing the equations we see

$$\begin{cases} x_1 + 4x_2 - 2x_3 = 4 \\ 2x_1 - 3x_2 + x_3 = -3 \end{cases}$$

$$\begin{cases} x_1 + 4x_2 - 2x_3 = 4 \\ \qquad 11x_2 - 5x_3 = 11 \end{cases}$$

$$\begin{cases} x_1 = 4 - 4x_2 + 2x_3 = \dfrac{2}{11}x_3 = \dfrac{2}{11}s \\ x_2 = 1 + \dfrac{5}{11}x_3 = 1 + \dfrac{5}{11}s \\ x_3 = s \in \mathbb{R}. \end{cases}$$

Solving Systems by Reducing Matrices

Recall for a system of equations, we can apply the following operations and the solution is unchanged:

1. interchange two equations,
2. multiply one equation by a nonzero constant, and
3. adding a constant multiple of an equation to another equation.

This is how we generally solve a system of equations.

Exercise. Solve

$$\begin{cases} x + 2y = 5 \\ 3x + 4y = 6. \end{cases}$$

Solution: Replace the first equation by the second equation minus 2 times the first equation:

$$\begin{cases} x = -4 \\ 3x + 4y = 6. \end{cases}$$

Then substitute $x = 4$ into the second equation:

$$\begin{cases} x = -4 \\ y = \dfrac{18}{4} = \dfrac{9}{2}. \end{cases}$$

Definition (Augmented Coefficient Matrix)**.** Given a matrix equation $AX = B$, the *augmented coefficient matrix* is the matrix $[A|B]$.

Remark. Recall that any system of linear equations can be written as a matrix equation. Thus, the augmented coefficient matrix is a matrix associated to such a system.

Example. Consider the system

$$\begin{cases} x + 2y = 5 \\ 3x + 4y = 6. \end{cases}$$

Associated to it is the matrix equation $AX = B$ with

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \qquad X = \begin{bmatrix} x \\ y \end{bmatrix} \qquad B = \begin{bmatrix} 5 \\ 6 \end{bmatrix}.$$

Recall here A is coefficient matrix.

The augmented coefficient matrix for this system is

$$[A|B] = \left[\begin{array}{cc|c} 1 & 2 & 5 \\ 3 & 4 & 6 \end{array}\right].$$

Example (General Example)**.** The system

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + a_{n2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

corresponds to the augmented coefficient matrix

$$[A|B] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right].$$

Definition (Elementary Row Operations)**.** The following are *elementary row operations* one can perform on matrices:

1. interchange two rows ($R_i \leftrightarrow R_j$),
2. multiply one row by a nonzero constant (kR_i), and
3. adding a constant multiple of one row to another row ($kR_i + R_j$; replaces row j but leaves row i unchanged).

If a matrix B can obtained from another matrix A by applying one or more elementary row operations, then A is *equivalent* to B.

Example.

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} 3 & 4 \\ 1 & 2 \end{bmatrix} \qquad \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \xrightarrow{4R_1} \begin{bmatrix} 4 & 8 \\ 3 & 4 \end{bmatrix} \qquad \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \xrightarrow{2R_1 + R_2} \begin{bmatrix} 1 & 2 \\ 5 & 8 \end{bmatrix}$$

Matrix Reduction

Definition. A *zero-row* of a matrix is a row that consists entirely of zeros.

A *nonzero-row* of a matrix a row that is not a zero-row, i.e. it consists of at least one nonzero entry.

The first non-zero entry in a nonzero-row is called the *leading entry.*

Example. Consider

$$\begin{bmatrix} \mathbf{1} & 2 & 3 & 4 \\ 0 & 0 & \mathbf{5} & 6 \\ 0 & 0 & 0 & 0 \\ 0 & \mathbf{7} & 0 & 9 \end{bmatrix}.$$

The bolded entries are the leading entries. Row 3 is a zero-row, all other rows are nonzero-rows.

Definition (Reduced Matrix)**.** A matrix is said to be a *reduced matrix* if all of the following are true:

1. All zero-rows are at the bottom of the matrix.
2. For each non-zero row, the leading entry is 1, and all other entries in the column of the leading entry are 0.
3. The leading entry in each row is (strictly) to the right of the leading entry in any row above it.

Example. Consider

$$A = \begin{bmatrix} \mathbf{1} & 2 & 3 & 4 \\ 0 & 0 & \mathbf{5} & 6 \\ 0 & 0 & 0 & 0 \\ 0 & \mathbf{7} & 0 & 9 \end{bmatrix} \qquad B = \begin{bmatrix} \mathbf{1} & 2 & 3 & 4 \\ 0 & \mathbf{7} & 0 & 9 \\ 0 & 0 & \mathbf{5} & 6 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

A is not reduced. B is slightly better (satisfy conditions 1 and 3), but fails condition 2, so it is not reduced either.

Lecture Date: May 30, 2013

Example. Consider

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \qquad C = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \qquad F = \begin{bmatrix} 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

A, C, and E are not reduced, but B, D, F are.

Example. In general, a reduced matrix looks like:

$$\begin{bmatrix} 1 & 0 & * & 0 & * & 0 & * & * \\ 0 & 1 & * & 0 & * & 0 & * & * \\ 0 & 0 & 0 & 1 & * & 0 & * & * \\ 0 & 0 & 0 & 0 & 0 & 1 & * & * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Theorem. *Each matrix is equivalent to exactly one reduced matrix.*

Remark. The theorem tells us we can use elementary row operations to reduce *any* matrix to a reduced one; moreover there is only one answer.

General Algorithm #1 to Reduce a Matrix

1. Move all zero-rows to the bottom.
2. Order the nonzero-rows so that the leading entry of any row occurs to the right or below all leading entries in the rows above.
3. Make the first leading entry a 1. Use it to change all leading entries below it to 0.
4. Repeat last step for all nonzero-rows until there is nothing more to do.

The result will be a reduced matrix as in the theorem.

General Algorithm #2 to Reduce a Matrix

1. Move all zero-rows to the bottom.
2. Order the nonzero-rows so that the leading entry of any row occurs to the right or below all leading entries in the rows above.
3. Starting from the top left leading entry, make all the other entries in the column zero (by adding/subtracting multiples of one row to another).

4. Repeat last step for all nonzero-rows until there is nothing more to do.
5. Make all the leading entries a 1 (by dividing).

The result will be a reduced matrix as in the theorem.

Remark. Algorithm #1 is better for computer programmers as they can just write one recursive function. Algorithm #2 is better for us when calculating by hand because it avoids fractions and it's generally faster.

Example. Reduce

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

Solution:

$$\begin{bmatrix} 2 & 1 & 0 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \\ 3 & 1 & 0 & 1 \end{bmatrix} \xrightarrow{R_3 \leftrightarrow R_4} \begin{bmatrix} 2 & 1 & 0 & 3 \\ 0 & 0 & 1 & 2 \\ 3 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{R_2 \leftrightarrow R_3} \begin{bmatrix} 2 & 1 & 0 & 3 \\ 3 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\xrightarrow{-3R_1 + 2R_2} \begin{bmatrix} 2 & 1 & 0 & 3 \\ 0 & -1 & 0 & -7 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{-3R_1 + R_2} \begin{bmatrix} 2 & 0 & 0 & -4 \\ 0 & -1 & 0 & -7 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\xrightarrow{\frac{1}{2}R_1, -R_2} \begin{bmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 7 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Theorem. *Each matrix is equivalent to exactly one reduced matrix. In other words, any matrix can be reduced to an unique reduced matrix using elementary row operations.*

Remark. The theorem tells us we can use row operations to solve matrix equations!

Example. Consider again the augmented coefficient matrix

$$\left[\begin{array}{cc|c} 1 & 2 & 5 \\ 3 & 4 & 6 \end{array}\right].$$

We apply elementary row operations to it to obtain a sequence of equivalent matrices:

$$\left[\begin{array}{cc|c} 1 & 2 & 5 \\ 3 & 4 & 6 \end{array}\right] \xrightarrow{-2R_1} \left[\begin{array}{cc|c} -2 & -4 & -10 \\ 3 & 4 & 6 \end{array}\right]$$

$$\xrightarrow{R_1 + R_2} \left[\begin{array}{cc|c} 1 & 0 & -4 \\ 3 & 4 & 6 \end{array}\right]$$

$$\xrightarrow{-3R_1 + R_2} \left[\begin{array}{cc|c} 1 & 0 & -4 \\ 0 & 4 & 18 \end{array}\right]$$

$$\xrightarrow{\frac{1}{4}R_2} \left[\begin{array}{cc|c} 1 & 0 & -4 \\ 0 & 1 & \frac{9}{2} \end{array}\right].$$

The last matrix is the augmented coefficient matrix for the system

$$\begin{cases} x = -4 \\ y = \dfrac{9}{2}. \end{cases}$$

Example. Solve

$$\begin{cases} 2x + y = 3 \\ \qquad\; z = 2 \\ 3x + y = 1 \end{cases}$$

using matrices.

Solution: The augmented coefficient matrix is

$$\left[\begin{array}{ccc|c} 2 & 1 & 0 & 3 \\ 0 & 0 & 1 & 2 \\ 3 & 1 & 0 & 1 \end{array}\right].$$

It reduces to (see a previous example)

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 7 \\ 0 & 0 & 1 & 2 \end{array}\right].$$

Therefore the solution is $(x, y, z) = (2, 7, 2)$.

Example. Solve

$$\begin{cases} x + 2y + 4z - 6 = 0 \\ 2z + y - 3 = 0 \\ x + y + 2z - 1 = 0 \end{cases}$$

using matrices.

Solution: First, rewrite the system of equations:

$$\begin{cases} x + 2y + 4z = 6 \\ y + 2z = 3 \\ x + y + 2z = 1. \end{cases}$$

We reduce the corresponding augmented coefficient matrix:

$$\left[\begin{array}{ccc|c} 1 & 2 & 4 & 6 \\ 0 & 1 & 2 & 3 \\ 1 & 1 & 2 & 1 \end{array}\right] \xrightarrow{-R_1+R_3} \left[\begin{array}{ccc|c} 1 & 2 & 4 & 6 \\ 0 & 1 & 2 & 3 \\ 0 & -1 & -2 & -5 \end{array}\right] \xrightarrow{R_2+R_3} \left[\begin{array}{ccc|c} 1 & 2 & 4 & 6 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & -2 \end{array}\right]$$

$$\xrightarrow{-2R_2+R_1} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & -2 \end{array}\right] \xrightarrow{-\frac{1}{2}R_3} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{-3R_3+R_2} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right].$$

This corresponds to the system

$$\begin{cases} x = 0 \\ y + 2z = 0 \\ 0 = 1. \end{cases}$$

The last equation is absurd! Therefore there is no solution.

Parametrized Family of Solutions

Example (One-Parameter Family of Solutions)**.** Solve

$$\begin{cases} x_1 + 4x_2 - 2x_3 = 4 \\ 2x_1 - 3x_2 + x_3 = -3 \end{cases}$$

using matrices.

Solution: We reduce the augmented coefficient matrix:

$$\left[\begin{array}{ccc|c} 1 & 4 & -2 & 4 \\ 2 & -3 & 1 & -3 \end{array}\right] \xrightarrow{2R_1 - R_2} \left[\begin{array}{ccc|c} 1 & 4 & -2 & 4 \\ 0 & 11 & -5 & 11 \end{array}\right]$$
$$\xrightarrow{\frac{1}{11}R_2} \left[\begin{array}{ccc|c} 1 & 4 & -2 & 4 \\ 0 & 1 & -\frac{5}{11} & 1 \end{array}\right]$$
$$\xrightarrow{-4R_2 + R_1} \left[\begin{array}{ccc|c} 1 & 0 & -\frac{2}{11} & 0 \\ 0 & 1 & -\frac{5}{11} & 1 \end{array}\right].$$

The associated system is

$$\begin{cases} x_1 - \dfrac{2}{11}x_3 = 0 \\ x_2 - \dfrac{5}{11}x_3 = 1. \end{cases}$$

Since we have more variables than equations, we make x_3 a parameter. The general solution is then

$$\begin{cases} x_1 = \dfrac{2}{11}s \\ x_2 = 1 + \dfrac{5}{11}s \\ x_3 = s \in \mathbb{R}. \end{cases}$$

Example (Two-Parameter Family of Solutions)**.** Solve

$$\begin{cases} x_1 - 2x_2 + 3x_3 = 4 \\ 2x_1 - 4x_2 + 6x_3 = 8 \\ -3x_1 + 6x_2 - 9x_3 = -12 \end{cases}$$

using matrices.

Solution: We reduce the augmented coefficient matrix:

$$\left[\begin{array}{ccc|c} 1 & -2 & 3 & 4 \\ 2 & -4 & 6 & 8 \\ -3 & 6 & -9 & -12 \end{array}\right] \xrightarrow{-2R_1 + R_2} \left[\begin{array}{ccc|c} 1 & -2 & 3 & 4 \\ 0 & 0 & 0 & 0 \\ -3 & 6 & -9 & -12 \end{array}\right] \xrightarrow{3R_1 + R_3} \left[\begin{array}{ccc|c} 1 & -2 & 3 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

The corresponding system is

$$\begin{cases} x_1 - 2x_2 + 3x_3 = 4 \\ 0 = 0 \\ 0 = 0. \end{cases}$$

The last two equations are redundant. Since we have more variables than equations, we make x_2 and x_3 parameters. The general solution is then

$$\begin{cases} x_1 = 4 + 2t - 3s \\ x_2 = t \in \mathbb{R} \\ x_3 = s \in \mathbb{R}. \end{cases}$$

Homogeneous and Nonhomogeneous Systems

Definition (Homogeneous and Nonhomogeneous Systems)**.** The system

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + a_{n2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

and the corresponding matrix equation

$$AX = B$$

and the corresponding augmented coefficient matrix

$$[A|B] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right].$$

If $B = 0$ (i.e. $b_1 = \cdots = b_m = 0$) then the system is called a *homogeneous system*. Otherwise, if $B \neq 0$ (i.e. some $b_i \neq 0$) then the system is called a *nonhomogeneous system*.

Remark. Therefore, a homogeneous system has matrix equation

$$AX = 0$$

and augmented coefficient matrix

$$[A|0].$$

Example. The system

$$\begin{cases} x + 2y + 4z = 6 \\ y + 2z = 3 \\ x + y + 2z = 1 \end{cases}$$

is nonhomogeneous.

The system

$$\begin{cases} x + 2y + 4z = 0 \\ y + 2z = 0 \\ x + y + 2z = 0 \end{cases}$$

is homogeneous.

Example. Recall how we solved the nonhomogeneous system

$$\begin{cases} x + 2y = 5 \\ 3x + 4y = 6 \end{cases}$$

by reducing the augmented coefficient matrix:

$$\left[\begin{array}{cc|c} 1 & 2 & 5 \\ 3 & 4 & 6 \end{array}\right] \xrightarrow{-2R_1} \left[\begin{array}{cc|c} -2 & -4 & -10 \\ 3 & 4 & 6 \end{array}\right] \xrightarrow{R_1+R_2} \left[\begin{array}{cc|c} 1 & 0 & -4 \\ 3 & 4 & 6 \end{array}\right]$$
$$\xrightarrow{-3R_1+R_2} \left[\begin{array}{cc|c} 1 & 0 & -4 \\ 0 & 4 & 18 \end{array}\right] \xrightarrow{\frac{1}{4}R_2} \left[\begin{array}{cc|c} 1 & 0 & -4 \\ 0 & 1 & \frac{9}{2} \end{array}\right].$$

To solve the homogeneous system

$$\begin{cases} x + 2y = 0 \\ 3x + 4y = 0 \end{cases}$$

we reduce the augmented matrix using the *exact same elementary row operations*:

$$\left[\begin{array}{cc|c} 1 & 2 & 0 \\ 3 & 4 & 0 \end{array}\right] \xrightarrow{-2R_1} \left[\begin{array}{cc|c} -2 & -4 & 0 \\ 3 & 4 & 0 \end{array}\right] \xrightarrow{R_1+R_2} \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 3 & 4 & 0 \end{array}\right]$$
$$\xrightarrow{-3R_1+R_2} \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 4 & 0 \end{array}\right] \xrightarrow{\frac{1}{4}R_2} \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \end{array}\right].$$

Notice this this case, the last column is always 0. We may as well ignore it and just reduce the coefficient matrix!

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

This corresponds to the system

$$\begin{cases} x = 0 \\ y = 0 \end{cases}$$

and the solution is $x = y = 0$.

Remark. A homogeneous system

$$AX = 0$$

always has a solution: $X = 0$. Indeed,

$$A0 = 0.$$

This solution is called the *trivial solution.*

Theorem. *Let A be the reduced coefficient matrix of a homogeneous system of m linear equations and n unknowns, and let k be the number of nonzero-rows. Then $k \leq n$. Moreover,*

1. *if $k < n$ then the system has infinitely many solutions;*
2. *if $k = n$ then the system has an unique solution (the trivial solution).*

If a system fewer equations then unknowns, then $m < n$. Then the number of zero-rows k must satisfy $k \leq m$. Hence,

$$k < n$$

and we are in case 1 of the theorem. To rephrase:

Corollary. *A homogeneous system of linear equations with fewer equations then unknowns has infinitely many solutions.*

Example. The system

$$\begin{cases} x_1 + 4x_2 - 2x_3 = 0 \\ 2x_1 - 3x_2 + x_3 = 0 \end{cases}$$

is homogeneous with more unknowns than equations. By the corollary, it has infinitely many solutions.

Example. Solve

$$\begin{cases} x + 2y + 4z = 0 \\ \qquad y + 2z = 0 \\ \quad x + y + 2z = 0 \end{cases}$$

using matrices.

Solution: We reduce the corresponding 3×3 coefficient matrix:

$$\begin{bmatrix} 1 & 2 & 4 \\ 0 & 1 & 2 \\ 1 & 1 & 2 \end{bmatrix} \xrightarrow{-R_1+R_3} \begin{bmatrix} 1 & 2 & 4 \\ 0 & 1 & 2 \\ 0 & -1 & -2 \end{bmatrix} \xrightarrow{R_2+R_3} \begin{bmatrix} 1 & 2 & 4 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{-2R_2+R_1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$
$$\xrightarrow{-\frac{1}{2}R_3} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{-3R_3+R_2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$$

By the theorem, there are only $k = 2$ nonzero-rows. Therefore there are infinitely many solutions.

The corresponding system is

$$\begin{cases} \qquad x = 0 \\ y + 2z = 0 \end{cases}$$

and the solution in parametric form is therefore

$$x = 0 \qquad\qquad y = -2s \qquad\qquad z = s$$

where $s \in \mathbb{R}$.

Example. Solve

$$\begin{cases} 3x + 4y = 0 \\ x - 2y = 0 \\ 2x + y = 0 \\ 2x + 3y = 0 \end{cases}$$

using matrices.

Solution: We reduce the corresponding 2×4 coefficient matrix:

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

There are $k = 2$ nonzero-rows. Therefore by the theorem, the unique solution is the trivial solution $x = y = 0$.

Indeed the corresponding system is

$$\begin{cases} x = 0 \\ y = 0. \end{cases}$$

Relationship between homogeneous and nonhomogeneous system

For any matrix equation (or equivalently, a linear system) $AX = B$, we can form the homogeneous counterpart $AX = 0$.

We have two facts:

1. Suppose C is a solution to $AX = B$ (i.e. $AC = B$) and D is a solution to $AX = 0$ (i.e. $AD = 0$). Then $C + D$ is also a solution to $AX = B$ since
$$A(C + D) = AC + AD = B + 0 = B.$$
2. Suppose C' is another solution to $AX = B$. Then $C - C'$ is a solution to $AX = 0$ since
$$A(C - C') = AC - AC' = B - B = 0.$$

In other words:

1. A solution of $AX = B$ + a solution to $AX = 0$ is another solution to $AX = B$.
2. Any two solutions to $AX = B$ differs by a solution to $AX = 0$.

Putting it together and applying the previous theorem, we have the following:

Corollary. *If the nonhomogeneous system $AX = B$ has a solution, then it has the same number of solutions as the corresponding homogeneous system $AX = 0$.*

Precisely, let A be the reduced coefficient matrix of a nonhomogeneous system of m linear equations and n unknowns, and let k be the number of nonzero-rows. Then $k \leq n$, and

1. *if $k < n$ then the system has infinitely many solutions;*

2. *if $k = n$ then the system has an unique solution.*

Remark. Of course, unlike the homogeneous case, it is possible for the nonhomogeneous system $AX = B$ to have no solution. That is why the corollary only holds if it has a solution.

Lecture Date: June 4, 2013

Inverses

Example. Solve the linear equation

$$3x = 5.$$

Solution: We multiply both sides of the equation by $3^{-1} = \frac{1}{3}$:

$$\begin{aligned} 3^{-1}3x &= 3^{-1}5 \\ x &= 3^{-1}5 = \frac{5}{3}. \end{aligned}$$

In general, to solve the linear equation

$$ax = b$$

we multiply both sides of the equation by $a^{-1} = \frac{1}{a}$:

$$\begin{aligned} a^{-1}ax &= a^{-1}b \\ x &= a^{-1}b = \frac{b}{a}. \end{aligned}$$

This works because $a^{-1}a = \dfrac{a}{a} = 1$ (as long as $a \neq 0$).

We try this with a matrix equation

$$AX = B.$$

Suppose there is matrix C such that $CA = I$. Then we left multiply both sides of the equation by C:

$$\begin{aligned} CAX &= CB \\ IX = X &= CB. \end{aligned}$$

The solution is therefore $X = CB$. This only works if

1. such a matrix C exists, and
2. C is of the correct size so that CB makes sense.

Now, CB is a solution means

$$\begin{aligned} A(CB) &= B \\ (AC)B &= B, \end{aligned}$$

i.e. we also need $AC = I$. However, matrix multiplication is not commutative and in general $CA \neq AC$ (they may not even be the same size!). For example,

$$\begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \end{bmatrix} = I_1 \qquad \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \neq I_2$$

However, if A is square, then it can be shown that $CA = I$ implies $AC = I$ also.

Matrix Inverse

Definition (Matrix Inverse)**.** If A is a square matrix and there exists a matrix C such that $CA = I$, then C is called the *inverse of* A and A is said to be *invertible*.

Example.

1. If $A = \begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix}$ and $C = \begin{bmatrix} 7 & -2 \\ -3 & 1 \end{bmatrix}$ then $CA = \begin{bmatrix} 7 & -2 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.
2. If $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $C = \begin{bmatrix} -2 & 1 \\ 3/2 & -1/2 \end{bmatrix}$ then $CA = \begin{bmatrix} -2 & 1 \\ 3/2 & -1/2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

Theorem. *Suppose A is an invertible matrix. Then the inverse is unique. Moreover, if A^{-1} denotes the inverse, then*

$$A^{-1}A = I = AA^{-1}.$$

Exercise.

1. If $A = \begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix}$ and $A^{-1} = \begin{bmatrix} 7 & -2 \\ -3 & 1 \end{bmatrix}$ check that $A^{-1}A = I = AA^{-1}$.
2. If $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $A^{-1} = \begin{bmatrix} -2 & 1 \\ 3/2 & -1/2 \end{bmatrix}$ check that $A^{-1}A = I = AA^{-1}$.

Theorem. *Suppose A is an invertible matrix. Then A^{-1} is also invertible, and $\left(A^{-1}\right)^{-1} = A$.*

Theorem. *Suppose A and B are invertible matrices. Then $(AB)^{-1} = B^{-1}A^{-1}$ and $\left(A^{-1}\right)^T = \left(A^T\right)^{-1}$.*

Exercise. Simplify $\left(\left(A^T B^{-1} C\right)^{-1}\right)^T$.

Lecture Date: June 06, 2013

Solving Systems by Inverses

Theorem. *Suppose A is an invertible matrix. Then the matrix equation*

$$AX = B$$

has an unique solution $X = A^{-1}B$.

Remark. Recall that if a function f is one-to-one then it is invertible and the solution to $f(x) = b$ is $x = f^{-1}(b)$.

Remark. Since A is an invertible matrix, it is necessarily a square matrix, and therefore the number of equations = the number of unknowns.

Example. Solve

$$\begin{cases} x + 2y = 5 \\ 3x + 7y = 18. \end{cases}$$

Solution: The corresponding matrix equation is

$$AX = B$$
$$\begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5 \\ 18 \end{bmatrix}.$$

The coefficient matrix A is invertible, so the solution is

$$X = A^{-1}B = \begin{bmatrix} 7 & -2 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ 18 \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$$

.

Example. Solve

$$\begin{cases} x + 2y = 5 \\ 3x + 4y = 6. \end{cases}$$

Solution: The corresponding matrix equation is

$$AX = B$$
$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5 \\ 6 \end{bmatrix}.$$

The coefficient matrix A is invertible, so the solution is

$$X = A^{-1}B = \begin{bmatrix} -2 & 1 \\ 3/2 & -1/2 \end{bmatrix} \begin{bmatrix} 5 \\ 6 \end{bmatrix} = \begin{bmatrix} -4 \\ \frac{9}{2} \end{bmatrix}$$

.

Example. Show that $A = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$ is not invertible.

Solution: Suppose, conversely, that A is invertible. Write its inverse as $A^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Then

$$\begin{aligned} A^{-1}A &= I \\ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ \begin{bmatrix} 0 & a+b \\ 0 & c+d \end{bmatrix} &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \end{aligned}$$

This is absurd! Therefore, A is not invertible.

Finding Inverses via Elementary Row Operations

Theorem. *A square matrix A is invertible if and only if A is equivalent to I (i.e. there is a sequence of elementary row operations $E_1, E_2, \ldots, E_k$ that takes A to I).*

Moreover, the same sequence of elementary row operations takes I to A^{-1}.

The second part of the theorem thus tells us a way to find A^{-1}!

Algorithm to find the inverse of a matrix Let A be an $n \times n$ square matrix.

1. Form the $n \times 2n$ augmented matrix $[A|I]$.
2. Apply elementary row operations to row reduce $[A|I]$:
$$[A|I] \to \cdots \to [R|B].$$
3. If $R = I$, then A is invertible and $B = A^{-1}$; otherwise, A is not invertible.

Example. Find $\begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix}^{-1}$.

Solution: We form and reduce the augmented matrix:

$$\left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 7 & 0 & 1 \end{array}\right] \xrightarrow{-3R_1+R_2} \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & -3 & 1 \end{array}\right] \xrightarrow{-2R_2+R_1} \left[\begin{array}{cc|cc} 1 & 0 & 7 & -2 \\ 0 & 1 & -3 & 1 \end{array}\right].$$

Therefore, $\begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix}^{-1} = \begin{bmatrix} 7 & -2 \\ -3 & 1 \end{bmatrix}$.

Exercise. Find $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}^{-1}$ using the above algorithm.

Example. Solve

$$\begin{cases} x - 2z = 1 \\ 4x - 2y + z = 2 \\ x + 2y - 10z = -1. \end{cases}$$

Solution: The corresponding matrix equation is

$$AX = B$$
$$\begin{bmatrix} 1 & 0 & -2 \\ 4 & -2 & 1 \\ 1 & 2 & -10 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}.$$

We compute A^{-1}:

$$[A|I] = \left[\begin{array}{ccc|ccc} 1 & 0 & -2 & 1 & 0 & 0 \\ 4 & -2 & 1 & 0 & 1 & 0 \\ 1 & 2 & -10 & 0 & 0 & 1 \end{array}\right]$$
$$\xrightarrow{-4R_1+R_2, -R1+R_3} \left[\begin{array}{ccc|ccc} 1 & 0 & -2 & 1 & 0 & 0 \\ 0 & -2 & 9 & -4 & 1 & 0 \\ 0 & 2 & -8 & -1 & 0 & 1 \end{array}\right]$$
$$\xrightarrow{R_2+R_3} \left[\begin{array}{ccc|ccc} 1 & 0 & -2 & 1 & 0 & 0 \\ 0 & -2 & 9 & -4 & 1 & 0 \\ 0 & 0 & 1 & -5 & 1 & 1 \end{array}\right]$$
$$\xrightarrow{2R_3+R_1, -9R_3+R_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -9 & 2 & 2 \\ 0 & -2 & 0 & 41 & -8 & -9 \\ 0 & 0 & 1 & -5 & 1 & 1 \end{array}\right]$$
$$\xrightarrow{-\frac{1}{2}R_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -9 & 2 & 2 \\ 0 & 1 & 0 & -\frac{41}{2} & 4 & \frac{9}{2} \\ 0 & 0 & 1 & -5 & 1 & 1 \end{array}\right].$$

Therefore, the solution is

$$X = A^{-1}B$$
$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -9 & 2 & 2 \\ -\frac{41}{2} & 4 & \frac{9}{2} \\ -5 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} -7 \\ -17 \\ -4 \end{bmatrix}.$$

Example. Solve

$$\begin{cases} x + 2y + 4z = 0 \\ y + 2z = 0 \\ x + y + 2z = 0. \end{cases}$$

Solution: We try to find the inverse of the coefficient matrix:

$$[A|I] = \left[\begin{array}{ccc|ccc} 1 & 2 & 4 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 1 & 1 & 2 & 0 & 0 & 1 \end{array}\right] \to \cdots \to \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -2 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 & 1 \end{array}\right].$$

Thus the coefficient matrix is not invertible, so we can't use inverses to solve. However, by a previous example, the solution was found (via reduction) to be

$$x = 0 \qquad y = -2s \qquad z = s$$

where $s \in \mathbb{R}$.

Lecture Date: June 11, 2013

Determinants

The determinant of a square matrix is a real number associated to a matrix. It will tell us whether a matrix is invertible or not.

References Most of the material from this section is sourced from

1. Nicholson, W. Keith, *Elementary linear algebra, with applications*, second edition,
2. O'Nan, Michael and Herbert, Enderton, *Linear algebra*, third edition.

Definition (Leibniz formula for determinant)**.** The *determinant* of a square matrix $A = [a_{ij}]$ of order n is

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{\sigma(i),i}$$

where S_n is the permutation group of n elements and $\operatorname{sgn} : S_n \to \{1, -1\}$ is the sign function. [1]

The definition above is good for theoretical purposes, but is hard to use to actually calculate $\det(A)$.

Notation Often, $\det(A)$ is written $\det A$ or $|A|$.

Properties of Determinants

Theorem (Invertibility and determinants)**.** *Let A be a square matrix. Then A is invertible if and only if* $\det A \neq 0$*.*

Theorem (Properties of determinants)**.** *Let A and B be square matrices of the same order. Then*

1. $\det(AB) = (\det A)(\det B)$,

2. $\det(A) = \det(A^T)$,

3. $\det(A^{-1}) = \dfrac{1}{\det A}$.

Laplace Expansion

We will devise a way to calculate the determinant "inductively". That is, we will define the determinant of a matrix of order n using the definition of the determinant of matrix of order $n - 1$.

[1] See `http://en.wikipedia.org/wiki/Leibniz_formula_for_determinants` for details.

Definition. The *determinant of a square matrix of order* 1 is

$$\det([a]) = a.$$

Definition. The *determinant of a square matrix of order* 2 is

$$\det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = ad - bc.$$

Example (Notation)**.**

$$\det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \det\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

Example.

$$\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = (1)(1) - (0)(0) = 1 \qquad \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = (1)(4) - (2)(3) = -2$$

$$\begin{vmatrix} 2 & 1 \\ 3 & 5 \end{vmatrix} = (2)(5) - (1)(3) = 7 \qquad \begin{vmatrix} 0 & 0 \\ 0 & 0 \end{vmatrix} = (0)(0) - (0)(0) = 0$$

Definition. The *determinant of a square matrix of order* 3 is

$$\det\left(\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}\right) = aei + bfg + cdh - afh - bdi - ceg.$$

Example.

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = (1)(5)(9) + (2)(6)(7) + (3)(4)(8) - (1)(6)(8) - (2)(4)(9) - (3)(5)(7) = 0$$

$$\begin{vmatrix} 3 & 2 & 1 \\ 1 & 4 & 2 \\ 0 & 1 & 2 \end{vmatrix} = \cdots = 15$$

$$\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = \cdots = 1$$

Example. We will compute the determinant of a matrix of order 3 using the determinants of matrices of order 2:

$$\begin{aligned} \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} &= aei + bfg + cdh - afh - bdi - ceg \\ &= aei - afh - bdi + bfg + cdh - ceg \\ &= a(ei - fh) - b(di - fg) + c(dh - eg) \\ &= a\begin{vmatrix} e & f \\ h & i \end{vmatrix} - b\begin{vmatrix} d & f \\ g & i \end{vmatrix} + c\begin{vmatrix} d & e \\ g & h \end{vmatrix}. \end{aligned}$$

Definition (Minor and Cofactor)**.** Suppose A is a square matrix of order n, and the determinant of any square matrix of order $n-1$ is known.

The (i,j)-*minor of* A, denoted $M_{ij}(A)$, is the determinant of the matrix (of order $n-1$) formed by deleting row i and column j of A. The (i,j)-*cofactor of* A, denoted $C_{ij}(A)$, is defined by

$$C_{ij}(A) = (-1)^{i+j} M_{ij}(A).$$

Clearly, $C_{ij}(A) = \pm M_{ij}(A)$. The number $(-1)^{i+j}$ is called the *sign* of the (i,j)-th position.

Example. Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$. Then

$$M_{11}(A) = \begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} = -3 \qquad C_{11}(A) = (-1)^{1+1} M_{11}(A) = -3$$

$$M_{12}(A) = \begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} = -6 \qquad C_{12}(A) = (-1)^{1+2} M_{12}(A) = 6$$

$$M_{13}(A) = \begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix} = -3 \qquad C_{13}(A) = (-1)^{1+3} M_{13}(A) = -3.$$

In general, the sign of a position can be determined from the following:

$$\begin{bmatrix} + & - & + & - & + & \cdots \\ - & + & - & + & - & \cdots \\ + & - & + & - & + & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Lecture Date: June 13, 2013

Theorem (Laplace Expansion). *The determinant of a square matrix A n can be computed by using Laplace expansion along any row or column. More precisely, let $A = [a_{ij}]$. Then:*

- *the expansion along row i is*

$$\det A = \sum_{j=1}^{n} a_{ij} C_{ij}(A) = a_{i1}C_{i1}(A) + a_{i2}C_{i2}(A) + \cdots + a_{in}C_{in}(A),$$

- *the expansion along column j is*

$$\det A = \sum_{i=1}^{n} a_{ij} C_{ij}(A) = a_{1j}C_{1j}(A) + a_{2j}C_{2j}(A) + \cdots + a_{nj}C_{nj}(A),$$

Remark. The more zeros a row or column has, the easier it is to expand along that!

Example. Using Laplace expansion for a matrix of order 3 and expanding along the first row, we have

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - b \begin{vmatrix} d & f \\ g & i \end{vmatrix} + c \begin{vmatrix} d & e \\ g & h \end{vmatrix}.$$

Example. Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$. Find $\det A$.

Solution: If we expand along the first row,

$$\det A = 1 \begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} - 2 \begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} + 3 \begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix} = 0.$$

If we expand along the first column,

$$\det A = 1 \begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} - 4 \begin{vmatrix} 2 & 3 \\ 8 & 9 \end{vmatrix} + 7 \begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix} = 0.$$

If we expand along the second column,

$$\det A = -2 \begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} + 5 \begin{vmatrix} 1 & 3 \\ 7 & 9 \end{vmatrix} - 8 \begin{vmatrix} 1 & 3 \\ 4 & 6 \end{vmatrix} = 0.$$

Elementary Row Operators and Determinants

Recall the (slightly modified) elementary row operations:

1. interchange two rows ($R_i \leftrightarrow R_j$),
2. multiply one row by a (nonzero) constant (kR_i), and

3. adding a constant multiple of one row to another row ($kR_i + R_j$).

Theorem (Effects of row operations on determinants)**.** *Suppose A is a square matrix of order n. If B is obtained from A by*

1. *interchanging two rows ($R_i \leftrightarrow R_j$) then* $\det B = -\det A$*;*
2. *multiplying one row by a constant k (kR_i) then* $\det B = k \det A$*;*
3. *adding a row to a constant multiple of another row ($R_i \to R_i + kR_j$) then* $\det B = \det A$.

Example. Let $A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & -3 & 8 \\ 1 & 0 & -5 \end{bmatrix}$.

1. If $B = \begin{bmatrix} 2 & -3 & 8 \\ 1 & 2 & -1 \\ 1 & 0 & -5 \end{bmatrix}$ (i.e. $R_1 \leftrightarrow R_2$), then $\det A = -\det B$, i.e.

$$\begin{vmatrix} 1 & 2 & -1 \\ 2 & -3 & 8 \\ 1 & 0 & -5 \end{vmatrix} = - \begin{vmatrix} 2 & -3 & 8 \\ 1 & 2 & -1 \\ 1 & 0 & -5 \end{vmatrix}$$

2. If $B = \begin{bmatrix} 3 & 6 & -3 \\ 2 & -3 & 8 \\ 1 & 0 & -5 \end{bmatrix}$ (i.e. $3 \times R_1$), then $3 \det A = \det B \Leftrightarrow \det A = \frac{1}{3} \det B$, i.e.

$$\begin{vmatrix} 1 & 2 & -1 \\ 2 & -3 & 8 \\ 1 & 0 & -5 \end{vmatrix} = \frac{1}{3} \begin{vmatrix} 3 & 6 & -3 \\ 2 & -3 & 8 \\ 1 & 0 & -5 \end{vmatrix}$$

3. If $B = \begin{bmatrix} 1 & 2 & -1 \\ 2 & -3 & 8 \\ 5 & -6 & 11 \end{bmatrix}$ (i.e. $R3 \to R_3 + 2R_2$), then $\det A = \det B$, i.e.

$$\begin{vmatrix} 1 & 2 & -1 \\ 2 & -3 & 8 \\ 1 & 0 & -5 \end{vmatrix} = \frac{1}{3} \begin{vmatrix} 1 & 2 & -1 \\ 2 & -3 & 8 \\ 5 & -6 & 11 \end{vmatrix}$$

Corollary. *If A is a square matrix of order n, then* $\det(kA) = k^n \det A$.

Corollary. *Suppose A is a square matrix. Then* $\det(A) = 0$ *if either*

1. *A has a zero-row,*
2. *A has two identical rows, or*
3. *one row of A is a scalar multiple of another.*

Recall $\det(A^T) = \det A$. Therefore by considering the transpose, row operations become column operations. Therefore, we can replace "row" by "column" in all the theorems and corollaries in this section.

Theorem (Effects of column operations on determinants)**.** *Suppose A is a square matrix of order n. If B is obtained from A by*

1. *interchanging two columns ($C_i \leftrightarrow C_j$) then* $\det B = -\det A$*;*
2. *multiplying one column by a constant k (kC_i) then* $\det B = k \det A$*;*
3. *adding a column to a constant multiple of another column ($C_i \to C_i + kC_j$) then* $\det B = \det A$*.*

Corollary. *Suppose A is a square matrix. Then* $\det(A) = 0$ *if either*

1. *A has a zero-column,*
2. *A has two identical column, or*
3. *one column of A is a scalar multiple of another.*

Exercise. Using row/column operations and the theorem above, show

1. $\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = 0.$
2. $\begin{vmatrix} 2 & 0 & 1 & -1 \\ 0 & 3 & -1 & 2 \\ 1 & 1 & 1 & 1 \\ -1 & 2 & 2 & 1 \end{vmatrix} = -20.$

Cramer's Rule

Recall one way to solve $AX = B$ is to use matrix inverse:

$$X = A^{-1}B.$$

Write $A = \begin{bmatrix} A_1 & A_2 & \cdots & A_n \end{bmatrix}$ where A_j is the j-th column of A.

Theorem. *The solution to $AX = B$, where $X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, is given by*

$$x_i = \frac{\det \begin{bmatrix} A_1 & \cdots & A_{i-1} & B & A_{i+1} & \cdots & A_n \end{bmatrix}}{\det A}.$$

The matrix in the numerator is A with the i-th column replaced by B.

Example. Recall the solution to

$$\begin{cases} x + 2y = 5 \\ 3x + 7y = 18. \end{cases}$$

is $x = -1$, $y = 3$.

We use Cramer's rule:

$$x = \frac{\begin{vmatrix} 5 & 2 \\ 18 & 7 \end{vmatrix}}{\begin{vmatrix} 1 & 2 \\ 3 & 7 \end{vmatrix}} = -1 \qquad y = \frac{\begin{vmatrix} 1 & 5 \\ 3 & 18 \end{vmatrix}}{\begin{vmatrix} 1 & 2 \\ 3 & 7 \end{vmatrix}} = 3.$$

Remark. Cramer's rule is useful when the coefficients of the linear system are so large or "ugly" such that the method of row reduction is hard. For example, solving

$$\begin{cases} 8464700x - 255871.22587z = 46 \\ 454.335874x - \sqrt{3}0.336587y + 86674z = 908.0336 \\ 845.35234x + 7574.111y - \sqrt{2324}z = -995.82829. \end{cases}$$

is hard if we tried to reduce the augmented coefficient matrix. However, Cramer's rule in this case is easier since all we have to compute are determinants.

Exercise. Use Cramer's rule to solve

$$\begin{cases} x - 2z = 1 \\ 4x - 2y + z = 2 \\ x + 2y - 10z = -1 \end{cases}$$

and verify the answer with a previous example.

Cofactor Matrix, Adjoint, and Matrix Inverse

Recall the definition of the (i, j)-th cofactor $C_{ij}(A)$.

Definition (Cofactor Matrix, Adjoint)**.** Let A be a square matrix. The *cofactor matrix of* A is defined to be $[C_{ij}(A)]$. The *adjoint of* A is defined to be,

$$\operatorname{adj}(A) = [C_{ij}(A)]^T$$

Example. The adjoint of $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is $\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

Recall $\det A \neq 0$ if and only if A is invertible.

Theorem. *If* $\det A \neq 0$ *then* $A^{-1} = \dfrac{1}{\det A}\operatorname{adj}(A)$.

Example.

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

Exercise. Show that

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \times \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Lecture Date: June 27, 2013

Functions of Several Variables

The Coordinate System

Any point P on the two-dimensional plane $\mathbb{R}^2$ can be represented by the *coordinates of* P, say $P = (x_0, y_0)$, also called an *ordered pair.*

Similarly any point Q (dimension 0) in the three-dimensional space $\mathbb{R}^3$ can be represented by the *coordinates of* Q, say

$$Q = (x_1, y_1, z_1),$$

also called an *ordered triple.* A line (through Q) has dimension 1 and can be represented parametrically by

$$\begin{cases} x = x_1 + as \\ t = y_1 + bs \\ z = z_1 + cs \end{cases} \quad \text{where } s \in \mathbb{R}.$$

A plane (through Q) has dimension 2 and can be represented parametrically by

$$\begin{cases} x = x_1 + ps + p't \\ y = y_1 + qs + q't \\ z = z_1 + rs + r't \end{cases} \quad \text{where } s, t \in \mathbb{R}$$

or by the equation

$$ax + by + cz = d.$$

In particular,

- the x, y-plane has equation $z = 0$,
- the x, z-plane has equation $y = 0$, and
- the y, z-plane has equation $x = 0$.

Cartesian Product

Definition (Cartesian Product)**.** Given two sets X and Y, we can form the *Cartesian product* $X \times Y$, defined by

$$X \times Y = \{(x, y) : x \in X, y \in Y\}.$$

The pair (x, y) is an *ordered pair.*

Example. $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$.

We can form the Cartesian product of a Cartesian product!

Example. $(X \times Y) \times Z = X \times Y \times Z = \{(x, y, z) : x \in X, y \in Y, z \in Z\}$. Here, (x, y, z) is an ordered triple.

Example.

- $\mathbb{R}^2 \times \mathbb{R} = \mathbb{R}^3$.
- $\mathbb{R}^2 \times \mathbb{R}^3 = \mathbb{R}^5$.
- $\mathbb{N}^3 \times \mathbb{N}^5 \times \mathbb{N}^7 = \mathbb{N}^1 5$.

Remark. $\dim(X \times Y) = \dim X + \dim Y$.

Example. If $X = \{0, 1, 2\}$ and $Y = \{0, 1\}$, then

$$X \times Y = \{(0,0), (1,0), (2,0), (1,0), (1,1), (1,2)\}$$

Definition (Cartesian Product)**.** Given sets $X_1, X_2, \ldots, X_n$, we can form the *Cartesian product* $X_1 \times X_2 \times \cdots X_n = \prod_{i=1}^{n} X_i$, defined by

$$\prod_{i=1}^{n} X_i = \{(x_1, x_2, \ldots, x_n) : x_i \in X_i \forall i = 1, \ldots, n\}.$$

An element $(x_1, x_2, \ldots, x_n)$ is an *n-tuple*.

Functions of One Variable

Given two sets X and Y, we can define a *function*

$$\begin{aligned} f : X &\longrightarrow Y \\ x &\longmapsto f(x) = y. \end{aligned}$$

We can view f as a rule that assigns to every (input) element $x \in X$ to one (output) element $y = f(x) \in Y$. The set X is the *domain of* f and the set Y is the *target of* f.

The *graph of the function* $f : X \to Y$ is a subset of $X \times Y$ defined by

$$\text{graph}(f) = \{(x, y) \in X \times Y : y = f(x)\}.$$

Example. If $X = \mathbb{R} = (-\infty, \infty)$ and $Y = \mathbb{R} = (-\infty, \infty)$, then

$$f : \mathbb{R} \to \mathbb{R}$$

is the familiar class of functions that takes a real number to a real number.

The graph of such functions is a curve in $\mathbb{R}^2$.

Example. If $X = \mathbb{N} = \{1, 2, 3, \ldots\}$ and $Y = \mathbb{R}$

$$f : \mathbb{N} \to \mathbb{R}$$

is called a *sequence*.

For example, $f(n) = n^2$ is the sequence $1, 4, 9, 16, \ldots$.

Functions of Several Variables

Similarly, given three sets X, Y, Z, we can define a *function*

$$\begin{aligned} f : X \times Y &\longrightarrow Z \\ (x, y) &\longmapsto f(x, y) = z. \end{aligned}$$

Again, we can view f as a rule that assigns to every (input) element pair $(x, y) \in X \times Y$ to one (output) element $z = f(x, y) \in Z$. The set $X \times Y$ is the *domain of* f and the set Z is the *target of* f.

The *graph of the function* $f : X \times Y \to Z$ is a subset of $X \times Y \times Z$ defined by

$$\operatorname{graph}(f) = \{(x, y, z) \in X \times Y \times Z : z = f(x, y)\}.$$

Example. If $X = Y = Z = \mathbb{R}$, then $X \times Y = \mathbb{R}^2$ and

$$f : \mathbb{R}^2 \to \mathbb{R}$$

is a function that takes a pair of real numbers and outputs a real number.

The graph of such a function is a surface in $\mathbb{R}^3$.

Definition (Function of Several Variables)**.** Given sets $X_1, X_2, \ldots, X_n$ and Y, a Y-valued *multivariate function* is

$$\begin{aligned} f : X_1 \times \cdots \times X_n &\longrightarrow Y \\ (x_1, \ldots, x_n) &\longmapsto f(x_1, \ldots, x_n) = y. \end{aligned}$$

The *graph of* f is a subset of $X_1 \times \cdots \times X_n \times Y$ given by

$$\operatorname{graph}(f) = \{(x_1, \ldots, x_n, y) \in X_1 \times \cdots \times X_n \times Y : f(x_1, \ldots, x_n) = y\}$$

Example.

1. $f : \mathbb{R}^2 \to \mathbb{R}$ defined by $f(x, y) = 2x + 3y$ is a function of two variables. For example, $f(1, 2) = 8$, $f(-3, 4) = 6$, etc.

2. $g : \mathbb{R}^2 \to \mathbb{R}$ defined by $g(x, y) = xy + 3x$ is a function of two variables. For example, $g(1, 2) = 5$, $g(-3, 4) = -21$, etc.

Example. Let $X = \{0, 1, 2\}$, $Y = \{0, 1\}$, and $Z = \mathbb{R}$. We define $h : X \times Y \to \mathbb{R}$ by

$$\begin{aligned} &h(0, 0) = -2 && h(1, 0) = 1/2 && h(2, 0) = \sqrt{123} \\ &h(0, 1) = \pi^2/5 && h(1, 1) = 0 && h(2, 1) = 3^{\sqrt[5]{2\pi}}. \end{aligned}$$

Example. The domain of the function defined by

$$f(x, y) = \frac{1}{x^2 + y^2}$$

is $\mathbb{R}^2 - \{(0, 0)\} = \big\{(x, y) \in \mathbb{R}^2 : (x, y) \neq (0, 0)\big\} = ((-\infty, 0) \cup (0, \infty)) \times ((-\infty, 0) \cup (0, \infty))$.

Example. The domain of the function defined by

$$g(x,y) = \frac{x+3}{y-2}$$

is $\mathbb{R}^2 - \mathbb{R} \times \{2\} = \left\{(x,y) \in \mathbb{R}^2 : y \neq 2\right\} = (-\infty, \infty) \times ((-\infty, 2) \cup (2, \infty))$.

Example. If

$$z^2 = x^2 + y^2$$

then $z \neq f(x,y)$ for any function f since, for example, $(5)^2 = (3)^2 + (4)^2$ and $(-5)^2 = (3)^2 + (4)^2$, and in general, $z = \pm\sqrt{x^2 + y^2}$ (two outputs).

Exercise. A company produces two products, X and Y. It costs \$20 to produce one unit of X and \$12 to produce one unit of Y, and sells them at \$35 and \$18 each respectively. Write down the

1. total cost function,
2. total revenue function, and
3. total profit function.

Sketching Planes and Surfaces

Exercise. Sketch the following surfaces:

1. $x + 2y + 5z = 10$.
2. $x + 5z = 10$.
3. $x + 2y = 10$.
4. $z = x^2$.
5. $z = f(x)$ where f is your favourite function.
6. $x^2 + y^2 + z^2 = r^2$.

Lecture Date: July 2, 2013

Sketching a Function of Two Variables using Level Curves

Recall if $f : \mathbb{R}^2 \to \mathbb{R}$ then

$$\text{graph}(f) = \left\{(x, y, z) \in \mathbb{R}^3 : z = f(x, y)\right\}.$$

Thus, we think of the function as $z = f(x, y)$.

Definition (Level curve)**.** If we fix $z = c$, then $f(x, y) = c$ is the one-dimensional *level curve* on the plane $z = c$

Exercise. Sketch some level curves for $z = x^2 + y^2$.

Exercise. Sketch $z = x^2 + y^2$.

Multivariable Calculus

Recall the definition of the derivative of a function $f : \mathbb{R} \to \mathbb{R}$ at the point a:

$$f'(a) = \lim_{h\to 0} \frac{f(a + h) - f(a)}{h} = \lim_{x\to a} \frac{f(x) - f(a)}{x - a}.$$

This is a *point-wise* definition, and represents the slope of the tangent line at $x = a$.

If f is differentiable at every $x \in S \subseteq \mathbb{R}$, then the derivative of f is

$$f'(x) = \lim_{h\to 0} \frac{f(x + h) - f(x)}{h}.$$

Partial Derivatives

We now deal with functions of several variables.

Definition (Partial Derivative)**.** Let $f(x, y)$ be a function of two variables and $(a, b) \in \text{domain}(f)$.

The *partial derivative of f with respect to x at the point* (a, b) is

$$f_x(a, b) = \lim_{h\to 0} \frac{f(a + h, b) - f(a, b)}{h}.$$

This represents the slope of the tangent line in the x direction at the point (a, b).

The *partial derivative of f with respect to y at the point* (a, b) is

$$f_y(a, b) = \lim_{h\to 0} \frac{f(a, b + h) - f(a, b)}{h}.$$

This represents the slope of the tangent line in the y direction at the point (a, b).

Definition (Partial Derivative)**.** Let $f(x,y)$ be a function of two variables and suppose the partial derivatives of f exist at $(x,y) \in S \subseteq \mathbb{R}^2$. The *partial derivative of* f *with respect to* x is

$$\frac{\partial f}{\partial x} = f_x(x,y) = \lim_{h\to 0} \frac{f(x+h,y) - f(x,y)}{h}$$

and the *partial derivative of* f *with respect to* y is

$$\frac{\partial f}{\partial y} = f_y(x,y) = \lim_{h\to 0} \frac{f(x,y+h) - f(x,y)}{h}.$$

Remark. To find, say f_x, treat y as a constant and take the derivative with repect to x as usual.

Notation If $z = f(x,y)$, then

$$\frac{\partial z}{\partial x} = \frac{\partial f}{\partial x} = f_x(x,y) \qquad \frac{\partial z}{\partial y} = \frac{\partial f}{\partial y} = f_y(x,y)$$

and

$$\left.\frac{\partial z}{\partial x}\right|_{(a,b)} = \left.\frac{\partial f}{\partial x}\right|_{(a,b)} = f_x(a,b) \qquad \left.\frac{\partial z}{\partial y}\right|_{(a,b)} = \left.\frac{\partial f}{\partial y}\right|_{(a,b)} = f_y(a,b)$$

Example. Let $f(x,y) = 4x^2 + 5xy - x^2y$. Find $f_x(x,y)$, $f_y(x,y)$, $f_x(1,2)$, $f_y(1,2)$.

Solution: :

$$\begin{aligned} f_x(x,y) &= 8x + 5y - 2xy & f_y(x,y) &= 5x - x^2 \\ f_x(1,2) &= 14 & f_y(1,2) &= 4. \end{aligned}$$

Exercise. Let $f(x,y) = 3x^2 + 2x^4y - x^3y^2 + xy - 3y^3$. Find $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$, $\left.\frac{\partial f}{\partial x}\right|_{(1,0)}$, and $\left.\frac{\partial f}{\partial y}\right|_{(1,0)}$.

Exercise. Let $z = 3x^2e^{2xy} - \sqrt{x^2 - 2y} + x^4y^{1/2} - \frac{1}{x^2+y^2}$. Find $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$.

Lecture Date: July 4, 2013

Definition (Partial Derivative)**.** Let $f(x_1, \ldots, x_n)$ be a function of n variables and suppose the partial derivatives of f exist at $(x_1, \ldots, x_n) \in S \subseteq \mathbb{R}^n$. The *partial derivative of f with respect to x_i* is

$$\frac{\partial f}{\partial x_i} = f_{x_i}(x, y) = \lim_{h \to 0} \frac{f(x_1, \ldots, x_i + h, \ldots, x_n) - f(x_1, \ldots, x_n)}{h}$$

for $i = 1, \ldots, n$.

Example. If $f(x, y, z, w) = x^2 y + 2yw^2 z$ then

$$\frac{\partial f}{\partial x} = 2xy \qquad \frac{\partial f}{\partial y} = x^2 + 2w^2 z \qquad \frac{\partial f}{\partial z} = 4ywz \qquad \frac{\partial f}{\partial w} = 2yw^2.$$

Applications of Partial Derivatives

Set $z = f(x, y)$. Recall the definition

$$\frac{\partial z}{\partial x} = \lim_{h \to 0} \frac{f(x+h, y)}{h} \qquad \text{and} \qquad \frac{\partial z}{\partial y} = \lim_{h \to 0} \frac{f(x, y+h)}{h}.$$

Thus

"*$\frac{\partial z}{\partial x}$ is the change in z with respect to x with y held fixed*"

and

"*$\frac{\partial z}{\partial y}$ is the change in z with respect to y with x held fixed*"

Example (Marginal Cost and Marginal Revenue)**.** Suppose a manufacturer produces x units of product X and y units of Y. The total cost is given by the *joint-cost function*

$$C = f(x, y).$$

In this case $\dfrac{\partial C}{\partial x}$ is the *(partial) marginal cost with respect to x* and $\dfrac{\partial C}{\partial y}$ is the *(partial) marginal cost with respect to y*.

Suppose a manufacturer sells x units of product X and y units of Y. The total revenue is given by the *joint-revenue function*

$$R = f(x, y).$$

In this case $\dfrac{\partial R}{\partial x}$ is the *(partial) marginal revenue with respect to x* and $\dfrac{\partial R}{\partial y}$ is the *(partial) marginal revenue with respect to y*.

Example (Marginal Cost)**.** Suppose the joint-cost function for a manufacturer is

$$C(x, y) = 2x^2 - 320x + 60\sqrt{y} - 15y + 123.$$

Find the marginal cost when $x = 100$ and $y = 225$.

Solution: : The marginal ocst with respect to x is

$$\frac{\partial C}{\partial x} = 4x - 320 \qquad \frac{\partial C}{\partial y} = \frac{30}{\sqrt{y}} - 15$$

$$\left.\frac{\partial C}{\partial x}\right|_{(100,225)} = 4(100) - 320 = 80 \qquad \left.\frac{\partial C}{\partial y}\right|_{(100,225)} = \frac{30}{\sqrt{225}} - 15 = -13.$$

Example (Marginal Productivity)**.** Suppose a manufacturer employs w works and r robots. The total productivity is given by the *production function*

$$P = f(x, y).$$

In this case $\frac{\partial P}{\partial w}$ is the *(partial) marginal productivity with respect to* w and $\frac{\partial P}{\partial r}$ is the *(partial) marginal productivity with respect to* r

Example (Competitive and Complementary Products)**.** Suppose the demand for products A and B in terms of their prices are given by

$$q_A = f(p_A, p_B) \qquad \text{and} \qquad q_B = g(p_A, p_B).$$

Then

- $\frac{\partial q_A}{\partial p_A}$ is the marginal demand for A with respect to p_A,
- $\frac{\partial q_A}{\partial p_B}$ is the marginal demand for A with respect to p_B,
- $\frac{\partial q_B}{\partial p_A}$ is the marginal demand for B with respect to p_B,
- $\frac{\partial q_B}{\partial p_B}$ is the marginal demand for B with respect to p_B.

Normally, $\frac{\partial q_A}{\partial p_A} < 0$ and $\frac{\partial q_B}{\partial p_B} < 0$.

1. If $\frac{\partial q_A}{\partial p_B} > 0$ and $\frac{\partial q_B}{\partial p_A} > 0$ then A and B are *competitive products* or *substitutes.*
2. If $\frac{\partial q_A}{\partial p_B} < 0$ and $\frac{\partial q_B}{\partial p_A} < 0$ then A and B are *complementary products.*

Example (Competitive and Complementary Products)**.** Suppose the demand for products A and B in terms of their prices are given by

$$q_A = -5p_A + 20p_B \qquad \text{and} \qquad q_B = 3p_A^2\sqrt{p_B}.$$

Then

$$\frac{\partial q_A}{\partial p_B} = 20 > 0 \qquad \frac{\partial q_B}{\partial p_A} = 6p_Ap_B > 0$$

Thus A and B are *competitive products* or *substitutes.*

Implicit Partial Differentiation

Implicit differential deals with situations where z is not defined as a function of x and y, but instead defined *implicitly* via an equation.

Example. Let $3x + 2y + z = 0$. Find $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$.

Solution: We take $\frac{\partial}{\partial x}$ of both sides:

$$\begin{aligned} 3 + \frac{\partial z}{\partial x} &= 0 \\ \frac{\partial z}{\partial x} &= -3. \end{aligned}$$

Similarly, if we take $\frac{\partial}{\partial y}$ of both sides:

$$\begin{aligned} 2 + \frac{\partial z}{\partial y} &= 0 \\ \frac{\partial z}{\partial y} &= -2. \end{aligned}$$

Indeed, if we rearrange for z:

$$z = -3x - 2y,$$

then

$$\frac{\partial z}{\partial x} = -3 \qquad \frac{\partial z}{\partial y} = -2.$$

Example. Let $x^2 + y^2 + z^2 = r^2$ where $r \in \mathbb{R}$ is a constant. Find $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$.

Solution: We take $\frac{\partial}{\partial x}$ of both sides:

$$\begin{aligned} 2x + 2z\frac{\partial z}{\partial x} &= 0 \\ \frac{\partial z}{\partial x} &= -\frac{x}{z}. \end{aligned}$$

Similarly, if we take $\frac{\partial}{\partial y}$ of both sides:

$$\begin{aligned} 2y + 2z\frac{\partial z}{\partial y} &= 0 \\ \frac{\partial z}{\partial y} &= -\frac{y}{z}. \end{aligned}$$

Example. Let $x^2 + y^2 + z^2 = r^2$ where $r \in \mathbb{R}$ is a constant. Find $\dfrac{\partial z}{\partial x}$ and $\dfrac{\partial z}{\partial y}$.

Solution: We take $\dfrac{\partial}{\partial x}$ of both sides:

$$\begin{aligned} 2x + 2z\frac{\partial z}{\partial x} &= 0 \\ \frac{\partial z}{\partial x} &= -\frac{x}{z}. \end{aligned}$$

Similarly, if we take $\dfrac{\partial}{\partial y}$ of both sides:

$$\begin{aligned} 2y + 2z\frac{\partial z}{\partial y} &= 0 \\ \frac{\partial z}{\partial y} &= -\frac{y}{z}. \end{aligned}$$

Higher-Order Partial Derivatives

Let $z = f(x, y)$. Recalled that $f_x = \dfrac{\partial f}{\partial x}$ and $f_y = \dfrac{\partial f}{\partial y}$; each can be regarded as a function of two variables. Thus, we can differential again!

Definition (Second-Order Partial Derivatives)**.**

$$\begin{aligned} f_{xx} &= (f_x)_x & f_{xy} &= (f_x)_y \\ f_{yx} &= (f_y)_x & f_{yy} &= (f_y)_y \end{aligned}$$

or

$$\begin{aligned} \frac{\partial^2 f}{\partial x^2} &= \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) & \frac{\partial^2 f}{\partial y \partial x} &= \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) \\ \frac{\partial^2 f}{\partial x \partial y} &= \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) & \frac{\partial^2 f}{\partial y^2} &= \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) \end{aligned}$$

The derivatives f_{xy} and f_{yx} are called *mixed partial derivatives*.

We can similarly define third-order (or higher) partial derivatives:

$$f_{xxx} = \frac{\partial^3 f}{\partial x^3}, \quad f_{xxy} = \frac{\partial^3 f}{\partial y \partial x^2}, \quad \cdots \quad f_{yxy} = \frac{\partial^3 f}{\partial y \partial x \partial y}, \quad \cdots \quad f_{yyy} = \frac{\partial^3 f}{\partial y^3}.$$

Theorem (Schwarz' theorem)**.** *Suppose $f : \mathbb{R}^n \to \mathbb{R}$ has continuous second-order partial derivatives in a neighbourhood of $(a_1, \ldots, a_n)$, then*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a_1, \ldots, a_n) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a_1, \ldots, a_n) \quad \forall i, j = 1, \ldots, n.$$

Remark. In our (usual) case, $n = 2$, i.e. $f : \mathbb{R}^2 \to \mathbb{R}$, and f is "nice", so the mixed partial derivatives are equal everywhere:

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}.$$

Example. Let $f(x, y) = 4x^2 + 5xy - x^2y$. Find f_{xx}, f_{xy}, f_{yx} and f_{yy}.

Solution: Recall

$$f_x = 8x + 5y - 2xy \qquad\qquad f_y = 5x - x^2.$$

Thus,

$$\begin{aligned} f_{xx} &= 8 - 2y & \qquad f_{xy} &= 5 - 2x \\ f_{yx} &= 5 - 2x & \qquad f_{yy} &= 0. \end{aligned}$$

Example. Let $f(x, y) = 4x^2 + 5xy - x^2y$. Find f_{xxx}, f_{yxx}, f_{xxy}, and f_{yyx}.

Solution: Recall

$$f_{xx} = 8 - 2y \qquad f_{xy} = f_{yx} = 5 - 2x \qquad f_{yy} = 0.$$

Thus

$$f_{xxx} = 0 \qquad f_{yxx} = -2 \qquad f_{xyx} = -2 \qquad f_{yyx} = 0.$$

Lecture Date: July 9, 2013

Exercise. Let $f(x, y, z, w) = x^2y + 2yw^2z$. Find f_{xxx}, f_{xyz}, f_{xwz}, f_{wyx}, f_{wxy}, etc.

Exercise (Non-equal mixed partials)**.** Let

$$f(x, y) = \begin{cases} \frac{xy(x^2-y^2)}{x^2+y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}.$$

Show $f_{xy}(0, 0) \neq f_{yx}(0, 0) = 1$ by showing $f_{xy}(0, 0) = -1$ but $f_{yx}(0, 0) = 1$.

Lecture Date: July 11, 2013

Example. Let $x^2 + y^2 + z^2 = r^2$ where $r \in \mathbb{R}$ is a constant. Find $\dfrac{\partial^2 z}{\partial x^2}$ and $\dfrac{\partial^2 z}{\partial y \partial x}$.

Solution: We take $\dfrac{\partial}{\partial x}$ of both sides:

$$\begin{aligned} 2x + 2z\frac{\partial z}{\partial x} &= 0 \\ \frac{\partial z}{\partial x} &= -\frac{x}{z}, \end{aligned}$$

and take $\dfrac{\partial}{\partial x}$ of both sides once more:

$$\begin{aligned} \frac{\partial^2 z}{\partial x^2} &= -\frac{z - x\frac{\partial z}{\partial x}}{z^2} \\ &= -\frac{z + \frac{x^2}{z}}{z^2} \\ &= -\frac{z^2 + x^2}{z^3}. \end{aligned}$$

For $\dfrac{\partial^2 z}{\partial y \partial x}$:

$$\begin{aligned} \frac{\partial}{\partial y}\left[\frac{\partial z}{\partial x}\right] &= \frac{\partial}{\partial y}\left[-\frac{x}{z}\right] \\ \frac{\partial^2 z}{\partial y \partial x} &= \frac{x}{z^2}\frac{\partial z}{\partial y} \\ &= \frac{x}{z^2}\left(-\frac{y}{z}\right) \\ &= -\frac{xy}{z^3} \end{aligned}$$

Chain Rule

Recall the chain rule for a one-variable function: Let $y = f(u)$ and $u = u(x)$. Then

$$\frac{dy}{dx} = \frac{df}{dx} = \frac{df}{du}\frac{du}{dx}.$$

Theorem (Chain Rule)**.** *Let $z = f(x,y)$ where $x = x(r,s)$ and $y = y(r,s)$. If f, x, and y have continuous partial derivatives, then z is a function of r and s and*

$$\frac{\partial z}{\partial r} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial r} \qquad \frac{\partial z}{\partial s} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s}.$$

Remark. In this case, x and y can be thought of as intermediate variables and

$$z = f(x,y) = f(x(r,s), y(r,s)) = z(r,s).$$

Example. Let $z = f(x,y) = x^3y + 2xy$ where $x = 3r + s^2$ and $y = e^{rs} - \sqrt{s}$. Find $\dfrac{\partial z}{\partial r}$ and $\dfrac{\partial z}{\partial s}$.

Solution: First,

$$\frac{\partial f}{\partial x} = 3x^2 + 2y \qquad \frac{\partial f}{\partial y} = x^3 + 2x$$

and

$$\begin{aligned} \frac{\partial x}{\partial r} &= 3 & \frac{\partial x}{\partial s} &= 2s \\ \frac{\partial y}{\partial r} &= se^{rs} & \frac{\partial y}{\partial s} &= re^{rs} - \frac{1}{\sqrt{s}}. \end{aligned}$$

Thus,

$$\begin{aligned} \frac{\partial z}{\partial r} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial r} \\ &= \left(3x^2 + 2y\right)(3) + \left(x^3 + 2x\right)\left(se^{rs}\right) \\ &= 3\left(3(3r + s^2)^2 + 2(e^{rs} - \sqrt{s})\right) + se^{rs}\left((3r + s^2)^3 + 2(3r + s^2)\right) \\ &= \cdots \end{aligned}$$

and

$$\begin{aligned} \frac{\partial z}{\partial s} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s} \\ &= \left(3x^2 + 2y\right)(2s) + \left(x^3 + 2x\right)\left(re^{rs} - \frac{1}{\sqrt{s}}\right) \\ &= 2s\left(3(3r + s^2)^2 + 2(e^{rs} - \sqrt{s})\right) + \left((3r + s^2)^3 + 2(3r + s^2)\right)\left(re^{rs} - \frac{1}{\sqrt{s}}\right) \\ &= \cdots. \end{aligned}$$

Example. Suppose a manufacturer produces products A and B. The total cost C of producing q_A units of A and q_B units of B is given by the joint-cost function

$$C(q_A, q_B) = 10q_A + 2q_Aq_B + 12q_B + 5000.$$

The demand functions for A and B are given by

$$q_A = \frac{100}{p_A\sqrt{p_B}} \qquad q_B = 100 - 5p_A - 2p_B$$

where p_A and p_B are the prices of A and B.

Find the rate of change of total cost with respect to the price of A when $p_A = 10$ and $p_B = 25$.

Solution: We need to find $\dfrac{\partial C}{\partial p_A}$:

$$\begin{aligned}\frac{\partial C}{\partial p_A} &= \frac{\partial C}{\partial q_A}\frac{\partial q_A}{\partial p_A} + \frac{\partial C}{\partial q_B}\frac{\partial q_B}{\partial p_A} \\ &= (10 + 2q_B)\left(\frac{-100}{p_A^2\sqrt{p_B}}\right) + (2q_A + 12)(-5) \\ \left.\frac{\partial C}{\partial p_A}\right|_{(p_A,p_B)=(10,25)} &= (10 + 2(0))\left(\frac{-100}{(10)^2\sqrt{(25)}}\right) + (2(2) + 12)(-5) \\ &= -82\end{aligned}$$

since $q_A = 2$ and $q_B = 0$ when $(p_A, p_B) = (10, 25)$.

Theorem (Chain Rule). *Let $z = f(x_1, \ldots, x_n)$ where $x_i = x_i(r_1, \ldots, r_m), i = 1, \ldots, n$. If f and all x_i's have continuous partial derivatives, then z is a function of $r_1, \ldots, r_m$ and*

$$\frac{\partial z}{\partial r_j} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\frac{\partial x_i}{\partial r_j} = \frac{\partial f}{\partial x_1}\frac{\partial x_i}{\partial r_j} + \cdots + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial r_j} \quad \forall j = 1, \ldots, m.$$

Example. If $z = f(v, w, x, y)$ and v, w, x, y are functions of r, s, t, then z is a function of r, s, t and

$$\begin{aligned}\frac{\partial z}{\partial r} &= \frac{\partial z}{\partial v}\frac{\partial v}{\partial r} + \frac{\partial z}{\partial w}\frac{\partial w}{\partial r} + \frac{\partial z}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial r} \\ \frac{\partial z}{\partial s} &= \frac{\partial z}{\partial v}\frac{\partial v}{\partial s} + \frac{\partial z}{\partial w}\frac{\partial w}{\partial s} + \frac{\partial z}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial s} \\ \frac{\partial z}{\partial t} &= \frac{\partial z}{\partial v}\frac{\partial v}{\partial t} + \frac{\partial z}{\partial w}\frac{\partial w}{\partial t} + \frac{\partial z}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial t}\end{aligned}$$

Exercise. Let

$$w = e^{x+y+z}(x^2 + y^2 + z^2)$$

where

$$x = (r - s)^2 \qquad y = (r + s)^2 \qquad z = \ln(r^2 + s).$$

Find $\dfrac{\partial w}{\partial r}$ and $\dfrac{\partial w}{\partial s}$.

Example. If $z = f(v, w, x, y)$ and v, w, x, y are functions of t, then z is a function of t (single-variable function) and

$$\frac{dz}{dt} = \frac{\partial z}{\partial v}\frac{dv}{dt} + \frac{\partial z}{\partial w}\frac{dw}{dt} + \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}.$$

Exercise. Let

$$w = x^2yz + xy^2z + xyz^2$$

and

$$x = e^t \qquad y = te^t \qquad z = t^2e^t.$$

Find $\dfrac{dw}{dt}$.

Exercise. Let $z = f(x, y)$ be defined implicitly by

$$x^2 + y^2 + z^2 = 9$$

and

$$x = rs^2 \qquad\qquad y = r + 5s.$$

Find $\dfrac{\partial z}{\partial r}$.

Lecture Date: July 16, 2013

Maxima and Minima

Recall for $f : \mathbb{R} \to \mathbb{R}$ the definition of a point a being a relative/local maximum:

$$f(a) \geq f(x)$$

for all x near a.

We extend the definition to a function of two variables.

Definition (Relative/Local Maximum and Minimum)**.** A function $f : \mathbb{R}^2 \to \mathbb{R}$, say $z = f(x, y)$, has a *relative/local maximum* at the point (a, b) if

$$f(a, b) \geq f(x, y)$$

for every (x, y) in a neighbourhood of (a, b).

f has a *relative/local minimum* at the point (a, b) if

$$f(a, b) \leq f(x, y)$$

for every (x, y) in a neighbourhood of (a, b).

We can extend even further.

Definition (Relative/Local Maximum and Minimum)**.** A function $f : \mathbb{R}^n \to \mathbb{R}$, say $z = f(x_1, \ldots, x_n)$, has a *relative/local maximum* at the point $(a_1, \ldots, a_n)$ if

$$f(a_1, \ldots, a_n) \geq f(x_1, \ldots, x_n)$$

for every $(x_1, \ldots, x_n)$ in a neighbourhood of $(a_1, \ldots, a_n)$.

f has a *relative/local minimum* at the point $(a_1, \ldots, a_n)$ if

$$f(a_1, \ldots, a_n) \leq f(x_1, \ldots, x_n)$$

for every $(x_1, \ldots, x_n)$ in a neighbourhood of $(a_1, \ldots, a_n)$.

Critical Points, First Derivative Test

Recall that if a is a relative/local maximum/minimum of a nice $f : \mathbb{R} \to \mathbb{R}$ then a is a critical point, i.e. $f'(a) = 0$.

Definition (Critical Point)**.** A point (a, b) for which both $f_x(a, b) = 0$ and $f_y(a, b) = 0$ is called a *critical point for f*.

Theorem (First derivative test for functions of two variables)**.** *Suppose (a, b) is a relative/local maximum or minimum of $f : \mathbb{R}^2 \to \mathbb{R}$ and that f is "nice locally at (a, b)", i.e. f_x and f_y exist for all points close to (a, b). Then*

$$f_x(a, b) = f_y(a, b) = 0.$$

Remark. In other words, the relative/local maximum/minimum (a, b) is a solution to the system

$$\begin{cases} f_x(x,y) = 0 \\ f_y(x,y) = 0. \end{cases}$$

Thus, to look for the relative/local maximum/minimum, we look for the critical points of f: they are the candidates.

Exercise. Write the theorem and the definition of critical point for $f : \mathbb{R}^n \to \mathbb{R}$.

Exercise. Find the critical points of the following functions:

1. $f(x, y) = x^2 - 3y^2 - 8x - 9y + 3xy$.
2. $f(x, y, z) = 2x^2 + xy + y^2 + 100 - z(x + y - 200)$.
3. $f(x, y) = y - y^2 - 3x - 6x^2$.

Lecture Date: July 18, 2013

Second Derivative Test

Recall the second derivative test for $f : \mathbb{R} \to \mathbb{R}$. Suppose $f'(a) = 0$. Then

- if $f''(a) > 0$ then a is a relative/local minimum.
- if $f''(a) < 0$ then a is a relative/local maximum.
- if $f''(a) = 0$ then nothing can be said (f just has a horizontal tangent at a, maybe a is a point of inflection).

Theorem (Second derivative test for functions of two variables)**.** *Suppose (a, b) is a critical point of $f : \mathbb{R}^2 \to \mathbb{R}$ and f has continuous second order partial derivatives f_{xx}, f_{xy}, and f_{yy} at every point (x, y) near (a, b). Set*

$$D(x, y) = f_{xx}(x, y) f_{yy}(x, y) - (f_{xy}(x, y))^2 .$$

- *If $D(a, b) > 0$ and $f_{xx}(a, b) > 0$ then (a, b) is a relative/local minimum.*
- *If $D(a, b) > 0$ and $f_{xx}(a, b) < 0$ then (a, b) is a relative/local maximum.*
- *If $D(a, b) < 0$ then f has a saddle point at (a, b).*
- *If $D(a, b) = 0$ then no conclusions can be drawn about an extremum at (a, b).*

Remark. If $D(a, b) > 0$ then $f_{xx}(a, b)$ and $f_{yy}(a, b)$ *necessarily* has the same sign (exercise). Thus, you can test whether $f_{xx}(a, b) > 0$ or $f_{yy}(a, b) > 0$, whichever is easiest.

Exercise. When $D(a, b) > 0$, show that $f_{xx}(a, b)$ and $f_{yy}(a, b)$ have the same sign.

Remark. Let

$$H_f(x, y) = \begin{bmatrix} f_{xx}(x, y) & f_{xy}(x, y) \\ f_{yx}(x, y) & f_{yy}(x, y) \end{bmatrix} .$$

$H_f(a, b)$ is called the *Hessian matrix*. Note that

$$D(x, y) = \det (H_f(x, y)) .$$

If $D(a, b) = \det (H_f(a, b)) > 0$ then $H_f(x, y)$ is said to be *positive definite at (a,b)*.

Exercise. Use the second derivative test to find the relative/local extrema of $f(x, y) = x^3 + y^3 - xy$.

Exercise. Use the second derivative test to find the relative/local extrema of $f(x, y) = y^2 - x^2$.

Exercise. Find the relative/local extrema of $f(x, y) = x^4 + (x - y)^4$.

Applications

Example (Maximum output)**.** The production of a company is given by a production function

$$P = P(l, k) = 0.54l^2 - 0.02l^3 + 1.89k^2 - 0.09k^3$$

where l and k are amounts of labour and capital respectively. Find the values of l and k that maximizes production.

Solution: The critical points are the solution to the system:

$$\begin{aligned} \frac{\partial P}{\partial l} &= 0 & \qquad \frac{\partial P}{\partial k} &= 0 \\ 1.08l - 0.06l^2 &= 0 & 3.78k - 0.27k^2 &= 0 \\ 0.06l(18 - l) &= 0 & 0.27k(14 - k) &= 0 \\ l &= 0, 18 & k &= 0, 14. \end{aligned}$$

Therefore the four critical points are $(0, 0)$, $(0, 14)$, $(18, 0)$, $(18, 14)$.

Now,

$$D(l, k) = P_{ll}P_{kk} - P_{lk}^2 = (1.08 - 0.12l)(3.78 - 0.54k)$$

and

$$D(0, 0) > 0 \qquad D(0, 14) < 0 \qquad D(18, 0) < 0 \qquad D(18, 14) > 0.$$

Also,

$$P_{ll}(0, 0) > 0 \qquad\qquad P_{ll}(18, 14) < 0.$$

Therefore, maximum output is obtained when $l = 18$ and $k = 14$.

Example (Maximum profit)**.** A company produces two products, A and B. The production cost is constant at \$2 per unit of A and \$3 per unit of B. The quantities sold are given by the joint-demand function

$$q_A = 400(p_B - p_A) \qquad\qquad q_B = 400(9 + p_A - 2p_B)$$

where p_A and p_B are the prices of A and B respectively. Determine the prices of A and B that will maximize profit.

Solution: The profit function is

$$\begin{aligned} P = P(p_A, p_B) &= (\text{profit/unit of A}) \times (\text{units of A sold}) + (\text{profit/unit of B}) \times (\text{units of B sold}) \\ &= (p_A - 2)q_A + (p_B - 2)q_B \\ &= (p_A - 2)(400(p_B - p_A)) + (p_B - 2)(400(9 + p_A - 2p_B)). \end{aligned}$$

The critical points is the solution to the system:

$$\begin{aligned} \frac{\partial P}{\partial p_A} &= 0 & \qquad \frac{\partial P}{\partial p_B} &= 0 \\ -2p_A + 2p_B - 1 &= 0 & 2p_A - 4p_B + 13 &= 0 \\ p_A &= 5.5 & p_B &= 6. \end{aligned}$$

Now,

$$D(p_A, p_B) = P_{p_A p_A} P_{p_B p_B} - P_{p_A p_B}^2 = (-800)(-1600) - (800)^2 > 0$$

and of course $D(5.5, 6) > 0$. Also, $P_{p_A p_A}(5.5, 6) = -800 < 0$.

Therefore, maximum profit is obtained when $p_A = 5.5$ and $p_B = 6$.

Example (Maximum profit for a monopolist)**.** A monopolist, practicing price discrimination, sells one product in two different markets, say market A and market B. Let q_A and q_B be the number of units sold in markets A and B respectively, and let $p_A = f(q_A)$ and $p_B = f(q_B)$ be the demand functions. The revenue functions are therefore

$$r_A = q_A p_A = q_A f(q_A) \qquad\qquad r_B = q_B p_B = q_B g(q_B).$$

Suppose the product is produced at one plant, and the cost of producing $q = q_A + q_B$ units is $c = c(q)$. The monopolist's profit is

$$P = r_A + r_B - c.$$

To maximize P, we solve

$$\begin{aligned} \frac{\partial P}{\partial q_A} &= 0 & \frac{\partial P}{\partial q_B} &= 0 \\ \frac{dr_A}{dq_A} + 0 - \frac{\partial c}{\partial q_A} &= 0 & 0 + \frac{dr_B}{dq_B} - \frac{\partial c}{\partial q_B} &= 0 \\ \frac{dr_A}{dq_A} - \frac{dc}{dq}\frac{\partial q}{\partial q_A} &= 0 & \frac{dr_B}{dq_B} - \frac{dc}{dq}\frac{\partial q}{\partial q_B} &= 0 \\ \frac{dr_A}{dlq_A} - \frac{dc}{dq} &= 0 & \frac{dr_B}{dq_B} - \frac{dc}{dq} &= 0 \end{aligned}$$

since $q = q_A + q_B$ and $\dfrac{\partial q}{\partial q_A} = \dfrac{\partial q}{\partial q_B} = 1$. Therefore, we see that

$$\frac{dr_A}{dq_A} = \frac{dc}{dq} = \frac{dr_B}{dq_B}.$$

Now, $\dfrac{dr_A}{dq_A}$ and $\dfrac{dr_B}{dq_B}$ are marginal revenues and $\dfrac{dc}{dq}$ is marginal cost.

So, to maximize profit, the monopolist (loosely speaking) should charge prices such that the marginal revenues in both markets are equal, and equal to the marginal cost.

Lecture Date: July 23, 2013

Constrained Optimization: Lagrange Multipliers

Example. Find the relative/local extrema of

$$w = f(x, y, z) = x^2 + y^2 + z^2$$

subject to

$$x - y + 2z = 6.$$

Solution: We can get rid of the constraint by substitution make w a two variable function and optimize as if there are no constraints.

Indeed, $x = y - 2z + 6$ and

$$w = (y - 2z + 6)^2 + y^2 + z^2.$$

Then solving

$$\frac{\partial w}{\partial y} = 4y - 4z + 12 = 0 \qquad \frac{\partial w}{\partial z} = -4y + 10z - 24 = 0$$

gives $(y, z) = (-1, 2)$ as the critical point.

Since

$$\frac{\partial^2 w}{\partial y^2} = 4 \qquad \frac{\partial^2 w}{\partial z^2} = 10 \qquad \frac{\partial^2 w}{\partial z \partial y} = -4$$

we see that

$$D(-1, 2) = (4)(10) - (-4)^2 = 24 > 0$$

and by the second derivative test, w has a relative/local minimum at $(1, -1, 2)$.

We were lucky in the example that we were able to "get rid of the constraint". But we won't be lucky all the time. This is where the *method of Lagrange multipliers* come in handy.

Single Constraint

Theorem (Method of Lagrange multipliers – single constraint)**.** *Suppose we have a function $f(x_1, \ldots, x_n)$ subject to the constraint*

$$g_(x_1, \ldots, x_n) = 0.$$

Define $F : \mathbb{R}^{n+1} \to \mathbb{R}$ by

$$F(x_1, \ldots, x_n, \lambda) = f(x_1, \ldots, x_n) - \lambda g(x_1, \ldots, x_n).$$

If $(a_1, \ldots, a_n)$ is critical point of f subject to g, then there exist a value λ_0 such that $(a_1, \ldots, a_n, \lambda_0)$ is a critical point of F.

Definition (Lagrange multiplier)**.** The value λ_0 is called the *Lagrange multiplier.*

Remark. The method of Lagrange multipliers thus "converts" the constrained optimization problem (optimize f subject to $g_1, \ldots, g_M$) to an unconstrained optimization problem (optimize F).

Lecture Date: July 25, 2013

Example. Find the relative/local extrema of

$$w = f(x, y, z) = x^2 + y^2 + z^2$$

subject to

$$x - y + 2z = 6.$$

Solution: Using method of Lagrange multipliers, we write the constraint as

$$g(x, y, z) = x - y + 2z - 6 = 0$$

and define

$$\begin{aligned} F(x, y, z, \lambda) &= f(x, y, z) - \lambda g(x, y, z) \\ &= x^2 + y^2 + z^2 - \lambda(x - y + 2z - 6). \end{aligned}$$

Solving the system

$$\begin{cases} F_x = 2x - \lambda = 0 \\ F_y = 2y + \lambda = 0 \\ F_z = 2z - 2\lambda = 0 \\ F_\lambda = -x + y - 2z + 6 = 0 \end{cases}$$

yields the critical point $(x, y, z, \lambda) = (1, -1, 2, 2)$. Therefore $(x, y, z) = (1, -1, 2)$ is a critical point of f subject to g.

Remark. The method of Lagrange multipliers does not tell us whether the critical point correspond to a relative/local maximum or minimum.

In applications, one may be able to decide if the critical point obtain is a maximum or minimum by the context of the question.

Exercise.

1. Find the critical points for $f(x, y) = 3x - y + 6$ subject to $x^2 + y^2 = 4$.
2. Find the critical points for $f(x, y, z) = xyz$ where $xyz \neq 0$ subject to $x+2y+3z = 36$.

Example. Suppose a company has a order for 200 units of its product. It has two factories for production, say factory X and factory Y, and wishes to distribute the its production between them. The total cost function is given by

$$C = C(x, y) = 2x^2 + xy + y^2 + 200$$

where x and y are outputs of factory X and Y respectively. How should the company distribute its production?

Solution: We wish to minimize $C = C(x, y)$ subject to $x + y = 200$. We now solve this problem using the method of Lagrange multipliers.

Exercise. Finish the example above.

Exercise. Read §17.7 Example 4.

Multiple Constraints

Theorem (Method of Lagrange multipliers – multiple constraints)**.** *Suppose we have a function $f(x_1, \ldots, x_n)$ subject to the constraints*

$$g_1(x_1, \ldots, x_n) = 0$$
$$\vdots$$
$$g_M(x_1, \ldots, x_n) = 0.$$

Define $F : \mathbb{R}^{n+M} \to \mathbb{R}$ by

$$F(x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_M) = f(x_1, \ldots, x_n) - \sum_{k=1}^{M} \lambda_k g_k(x_1, \ldots, x_n).$$

If $(a_1, \ldots, a_n)$ is critical point of f subject to $g_1, \ldots, g_M$, then there exist values $\lambda_{0_1}, \ldots, \lambda_{0_M}$ such that $(a_1, \ldots, a_n, \lambda_{0_1}, \ldots, \lambda_{0_M})$ is a critical point of F.

Definition (Lagrange multiplier)**.** The values $\lambda_{0_1}, \ldots, \lambda_{0_M}$ are called the *Lagrange multipliers.*

Exercise. Find the critical points for $f(x, y, z) = xy + yz$ subject to $x^2 + y^2 = 8$ and $yz = 8$.

Lecture Date: July 30, 2013

Integrals

Recall for $f : \mathbb{R} \to \mathbb{R}$, the definite integral

$$\int_a^b f(x)dx$$

is the area (with sign) under f over the interval $[a, b]$.

More generally, if $f(x) \geq g(x)$, then

$$\int_a^b f(x) - g(x)dx$$

is the area between f and g over the interval $[a, b]$.

More precisely, a definitely integral is a Riemann sum:

$$\int_a^b f(x)dx = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i^*)\Delta x_i$$

where $\{x_0, x_1, \ldots, x_n\}$ is a regular partition of $[a, b]$, and $x_i^* \in [x_{i-1}, x_i]$ and $\Delta x_i = \frac{b-a}{n}$.

Definition (Definite integral)**.** Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a (integrable) function of two variables. Let $R \subset \mathbb{R}^2$ be a "nice" subset of the plane. The *definite integral*

$$\iint_R f(x, y)\ dA$$

is the volume (with sign) of the solid under f over the region R.

Rectangular Regions

Example. Let $R = [1, 2] \times [3, 4]$ and $f(x, y) = xy$. Then

$$\iint_R f(x, y)\ dA = \iint_{[1,2]\times[3,4]} xy\ dA$$

is the volume of the solid under the surface $z = xy$ over the square $[1, 2] \times [3, 4]$.

Definition (Riemann sum of a definite integral)**.** Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a (integrable) function of two variables and $R = [a, b] \times [c, d] \subset \mathbb{R}^2$ be a rectangular region. Let $P = \{x_0, \ldots, x_n\} \times \{y_0, \ldots, y_n\}$ be a regular partition of $[a, b] \times [c, d]$. The *definite integral* is defined as the *Riemann sum*

$$\iint_{[a,b]\times[c,d]} f(x, y)\ dA = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i^*, y_i^*)\Delta R_i$$

where $(x_i^*, y_i^*) \in [x_i - x_{i-1}, y_i - y_{i-1}]$ and $\Delta R_i = |[x_i - x_{i-1}, y_i - y_{i-1}]| = \frac{b-a}{n}\frac{d-c}{n}$.

Theorem (Fubini's theorem for rectangular regions)**.** *Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a (integrable) function of two variables and $R = [a, b] \times [c, d] \subset \mathbb{R}^2$ be a rectangular region. Then the definite integral can be computed as*

$$\iint\limits_R f(x, y)\ dA = \int_c^d \int_a^b f(x, y)\ dx\ dy = \int_a^b \int_c^d f(x, y)\ dy\ dx.$$

Remark. Fubini's theorem says two things:

1. The definite integral $\iint_R f(x, y)\ dA$ can be computed as an iterated integral, and
2. the iterated integrals $\int_c^d \int_a^b f(x, y)\ dx\ dy$ and $\int_a^b \int_c^d f(x, y)\ dy\ dx$ are equal.

Example. Let $R = [1, 2] \times [3, 4]$ and $f(x, y) = xy$. Then by Fubini's theorem

$$\begin{aligned}
\iint\limits_R f(x, y)\ dA &= \iint\limits_{[1,2]\times[3,4]} xy\ dA \\
&= \int_3^4 \int_1^2 xy\ dx\ dy = \int_3^4 \left[\frac{x^2 y}{2}\right]_1^2\ dy = \int_3^4 2y - \frac{y}{2}\ dy \\
&= \int_3^4 \frac{3y}{2}\ dy = \left[\frac{3y^2}{4}\right]_3^4 = 12 - \frac{27}{4} \\
&= \frac{21}{4}.
\end{aligned}$$

Exercise. Draw R and evaluate $\iint_R f(x, y)\ dA$ in two different ways:

1. $R = [0, 3] \times [0, 4]$, $f(x, y) = x$.
2. $R = [1, 2] \times [1, 3]$, $f(x, y) = x^2 - y$.
3. $R = [0, 1] \times [0, 2]$, $f(x, y) = x + y$.

Lecture Date: August 1, 2013

Non-rectangular Regions

Definition (Type I and II regions)**.** A region $R \in \mathbb{R}^2$ is a *type I region* if

$$R = \{(x, y) \in \mathbb{R}^2 : a \le x \le b, g_1(x) \le y \le g_2(x)\}$$

where g_1 and g_2 are continuous functions.

A region $R \in \mathbb{R}^2$ is a *type II region* if

$$R = \{(x, y) \in \mathbb{R}^2 : c \le y \le d, h_1(y) \le x \le h_2(y)\}$$

where h_1 and h_2 are continuous functions.

Exercise. Draw R:

1. $R = \{(x, y) \in \mathbb{R}^2 : -1 \le x \le 1, 0 \le y \le 1 - x\}$.
2. $R = \{(x, y) \in \mathbb{R}^2 : c \le y \le d, y^2 \le x \le 3y\}$.

Theorem (Fubini's theorem for non-rectangular regions)**.** *Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a (integrable) function of two variables and R be a type I or II region. Then*

$$\iint\limits_R f(x, y)\ dA = \int_a^b \int_{g_1(x)}^{g_2(x)} f(x, y)\ dy\ dx$$

if R is type I, and

$$\iint\limits_R f(x, y)\ dA = \int_c^d \int_{h_1(x)}^{h_2(x)} f(x, y)\ dx\ dy$$

if R is type II.

Example. Evaluate $\displaystyle\int_{-1}^{1} \int_0^{1-x} (2x + 1)\ dy\ dx$.

Solution:

$$\begin{aligned}\int_{-1}^{1} \int_0^{1-x} (2x + 1)\ dy\ dx &= \int_{-1}^{1} [(2x + 1)y]_0^{1-x}\ dx \\ &= \int_{-1}^{1} (2x + 1)(1 - x)\ dx = \int_{-1}^{1} -2x^2 + x + 1\ dx \\ &= \left[\frac{-2x^3}{3} + \frac{x^2}{2} + x\right]_{-1}^{1} \\ &= \left(\frac{-2}{3} + \frac{1}{2} + 1\right) - \left(\frac{2}{3} + \frac{1}{2} - 1\right) = \frac{2}{3}\end{aligned}$$

Example. Evaluate $\int_0^3 \int_{y^2}^{3y} 5x\ dx\ dy$.

Solution:

$$\begin{aligned}\int_0^3 \int_{y^2}^{3y} 5x\ dx\ dy &= \int_0^3 \left[\frac{5}{2}x^2\right]_{y^2}^{3y} dy \\ &= \frac{5}{2}\int_0^3 (9y^2 - y^4)\ dy \\ &= \frac{5}{2}\left[3y^3 - \frac{y^5}{5}\right]_0^3 \\ &= \frac{5}{2}\left(3(3)^3 - \frac{(3)^5}{5}\right) = \frac{5}{2}\left(\frac{162}{5}\right) = 81\end{aligned}$$

Exercise. Evaluate:

1. $\int_0^1 \int_{3x}^{x^2} 14x^2 y\ dy\ dx$.
2. $\int_0^1 \int_0^y e^{x+y}\ dx\ dy$.

Higher Dimensions

Theorem (Fubini's theorem)**.** *Let $f : \mathbb{R}^3 \to \mathbb{R}$ be a (integrable) function of three variables and $R \subset \mathbb{R}^3$ described by*

$$R = \left\{(x, y, z) \in \mathbb{R}^3 : a \le x \le b, g_1(x) \le y \le g_2(x), h_1(x, y) \le z \le h_2(x, y)\right\}.$$

Then

$$\iiint_R f(x, y, z)\ dV = \int_a^b \int_{g_1(x)}^{g_2(x)} \int_{h_1(x,y)}^{h_2(x,y)} f(x, y, z)\ dz\ dy\ dx.$$

Example. Compute $\int_0^1 \int_0^x \int_0^{x-y} x\ dz\ dy\ dx$.

$$\begin{aligned}\int_0^1 \int_0^x \int_0^{x-y} x\ dz\ dy\ dx &= \int_0^1 \int_0^x [xz]_0^{x-y}\ dy\ dx = \int_0^1 \int_0^x x^2 - xy\ dy\ dx \\ &= \int_0^1 \left[x^2 y - \frac{xy^2}{2}\right]_0^x dx = \int_0^1 x^3 - \frac{x^3}{2}\ dx = \int_0^1 \frac{x^3}{2}\ dx \\ &= \left[\frac{x^4}{8}\right]_0^1 = \frac{1}{8}\end{aligned}$$

Exercise. Evaluate:

1. $\int_0^1 \int_0^2 \int_0^{3-y} xy^2 z^3\ dx\ dy\ dz$.
2. $\int_0^1 \int_{x^2}^x \int_0^{xy}\ dz\ dy\ dx$.

Applications

In statistics, if X is a continuous random variable, then there is a *probability density function* f associated to it and satisfy:

1. $f(x) \geq 0$, and
2. $\int_{-\infty}^{\infty} f(x)\ dx = 1$.

Loosely speaking, $f(x) = \Pr(X = x)$ (not really, since $\Pr(X = x) = 0$ for a continuous random variable). The *cumulative density function* F is given by

$$F(x) = \int_{-\infty}^{x} f(t)\ dt.$$

The interpretation is $F(x) = \Pr(X \leq x)$.

Now, if X and Y are continuous random variables, then there is a *joint probability density function* f associated to it and satisfy:

1. $f(x, y) \geq 0$, and
2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\ dy\ dx = 1$.

Loosely speaking, $f(x, y) = \Pr(X = x, Y = y)$. The *cumulative density function* F is given by

$$F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(s, t)\ dt\ ds.$$

The interpretation is $F(x) = \Pr(X \leq x, Y \leq y)$. Also, there are the *marginal probability density functions*

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\ dy \qquad\qquad f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\ dx.$$

Example. If the joint probability density function for X and Y is $f(x, y)$, then

$$\Pr(a \leq X \leq b, c \leq Y \leq d) = \int_{c}^{d} \int_{a}^{b} f(x, y)\ dx\ dy.$$