# Applied Linear Algebra

## OTTO BRETSCHER

http://www.prenhall.com/bretscher

## Chapter 2
## Linear Transformation

Chia-Hui Chang
Email: chia@csie.ncu.edu.tw
National Central University, Taiwan

September 27, 2004

## 2.1 Introduction to Linear Transformations and Their Inverse

See Figure 1

Encryption of a coordinate $\vec{x} = \begin{bmatrix} 5 \\ 42 \end{bmatrix}$ to $\vec{y}$ by the following code:

$$\begin{aligned} y_1 &= \quad x_1 + \ 3x_2 = 131 \\ y_2 &= \ 2x_1 + \ 5x_2 = 220 \end{aligned}$$

At the headquarter, $\vec{y} = \begin{bmatrix} 131 \\ 220 \end{bmatrix}$ is received. We need to determine the actual $\vec{x}$ by solve the linear system.

$$A\vec{x} = \vec{b}$$

i.e. $$\begin{aligned} x_1 + \ 3x_2 &= 131 \\ 2x_1 + \ 5x_2 &= 220 \end{aligned}$$

If $\vec{y} = \begin{bmatrix} 133 \\ 223 \end{bmatrix}$ We need to solve it again by:

$$\begin{aligned} x_1 + \ & 3x_2 = 133 \\ 2x_1 + \ & 5x_2 = 223 \end{aligned}$$

For a general formula, we need solve the system

$$\begin{aligned} x_1 + \ & 3x_2 = y_1 \\ 2x_1 + \ & 5x_2 = y_2 \end{aligned}$$

for arbitrary constants $y_1$ and $y_2$.

For sender: $\vec{x} \to \vec{y}$ (encoding)

For receiver: $\vec{y} \to \vec{x}$ (decoding)

$$\left| \begin{array}{l} x_1 + \ 3x_2 = y_1 \\ 2x_1 + \ 5x_2 = \qquad y_2 \end{array} \right| \begin{array}{c} \longrightarrow \\ -2(I) \end{array}$$

$$\left| \begin{array}{ll} x_1 + 3x_2 & = y_1 \\ -x_2 & = -2y_1 + y_2 \end{array} \right| \begin{array}{c} \longrightarrow \\ \div(-1) \end{array}$$

$$\left| \begin{array}{ll} x_1 + 3x_2 & = y_1 \\ x_2 & = 2y_1 - y_2 \end{array} \right| \begin{array}{c} -3(II) \\ \longrightarrow \end{array}$$

$$\left| \begin{array}{lllll} x_1 & & = & -5y_1 & +3y_2 \\ & x_2 & = & 2y_1 & -y_2 \end{array} \right|$$

The decoding formula is:

$$\begin{array}{lll} x_1 = & -5y_1 & +3y_2 \\ x_2 = & 2y_1 & -y_2 \end{array}$$

or $\vec{x} = B\vec{y}$, where $B = \begin{bmatrix} -5 & 3 \\ 2 & -1 \end{bmatrix}$

**Definition.** *We say that the matrix $B$ is the* **inverse** *of the matrix $A$ and write $B = A^{-1}$.*

$$\vec{x} \underset{\vec{x}=B\vec{y},\, B=\begin{bmatrix} -5 & 3 \\ 2 & -1 \end{bmatrix}}{\overset{\vec{y}=A\vec{x},\, A=\begin{bmatrix} 1 & 3 \\ 2 & 5 \end{bmatrix}}{\rightleftarrows}} \vec{y}$$

The coding transformation is represented as

$$\underbrace{\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}}_{\vec{y}} = \begin{bmatrix} x_1 + & 3x_2 \\ 2x_1 + & 5x_2 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 3 \\ 2 & 5 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}_{\vec{x}}$$

or succinctly, as $\boxed{\vec{y} = A\vec{x}}$.

A transformation of the form $\vec{y} = A\vec{x}$ is called a **linear transformation**.

**Function**: Consider two sets $X$ and $Y$. A function $T : X \to Y$ is a rule that associates with each element $x \in X$ a unique element $y \in Y$.

The set $X$ is called the *domain* and $Y$ is called its *codomain*.

**Definition.** *A function $T$ from $R^n$ to $R^m$ is called a* **linear transformation** *if there is an $m \times n$ matrix $A$ such that*

$T(\vec{x}) = A\vec{x}$, *for all $\vec{x}$ in $R^n$.*

**Example.** *The linear transformation system*

$$\begin{aligned} y_1 &= 7x_1 + 3x_2 - 9x_3 + 8x_4 \\ y_2 &= 6x_1 + 2x_2 - 8x_3 + 7x_4 \\ y_3 &= 8x_1 + 4x_2 \qquad\quad + 7x_4 \end{aligned}$$

*(a function from $R^4$ to $R^3$) can be represented by the $3 \times 4$ matrix*

$$A = \begin{bmatrix} 7 & 3 & -9 & 8 \\ 6 & 2 & -8 & 7 \\ 8 & 4 & 0 & 7 \end{bmatrix}$$

**Example.** *The identity transformation system*

$$\begin{aligned} y_1 &= x_1 \\ y_2 &= x_2 \\ &\vdots \\ y_n &= x_n \end{aligned}$$

*(a linear transformation from $R^n$ to $R^n$ whose output equals its input) is represented by $n \times n$ matrix*

$$A = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

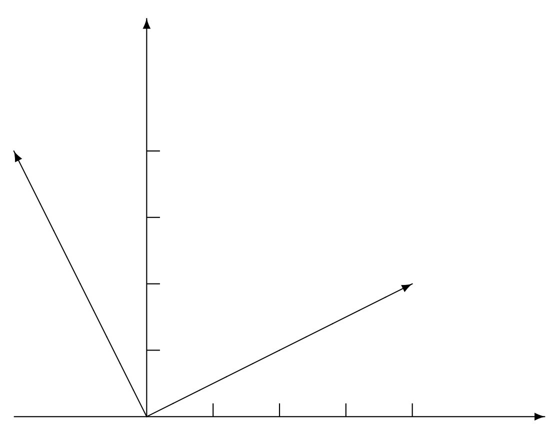
*This matrix is called the* **identiy matrix** *and is denoted by $I_n$:*

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},\ I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},\ etc.$$

**Example.** *Give a geometric interpretation of the linear transformation*

$$\vec{y} = A\vec{x}, \text{ where } A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix}.$$



See Figure 4 (pp.45).

**Fact 2.1.2** Consider a linear transformation $T$ from $R^n$ to $R^m$. Let $\vec{e}_i = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \leftarrow ith$

The matrix of $T$ can be represented as

$$A = \begin{bmatrix} | & | & & | \\ T(\vec{e}_1) & T(\vec{e}_2) & \cdots & T(\vec{e}_n) \\ | & | & & | \end{bmatrix}$$

Since

$$T(\vec{e}_i) = A\vec{e}_i = \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \vec{v}_i$$

**Example.** *Find the inverse for the following matrix:*

$$\begin{bmatrix} 1 & 2 \\ 3 & 9 \end{bmatrix}$$

**Solution**

$$\begin{bmatrix} 1 & 2 & \vdots & y_1 \\ 3 & 9 & \vdots & y_2 \end{bmatrix} \begin{matrix} \\ -3(I) \end{matrix}$$

$$\longrightarrow \begin{bmatrix} 1 & 2 & \vdots & y_1 \\ 0 & 3 & \vdots & -3y_1 + y_2 \end{bmatrix} \begin{matrix} \\ \div 3 \end{matrix}$$

$$\longrightarrow \begin{bmatrix} 1 & 2 & \vdots & y_1 \\ 0 & 1 & \vdots & -y_1 + \frac{1}{3}y_2 \end{bmatrix} \begin{matrix} -2(II) \\ \\ \end{matrix}$$

$$\longrightarrow \begin{bmatrix} 1 & 0 & \vdots & 3y_1 - \frac{2}{3}y_2 \\ 0 & 1 & \vdots & -y_1 + \frac{1}{3}y_2 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 1 & 2 \\ 3 & 9 \end{bmatrix}^{-1} = \begin{bmatrix} 3 & -\frac{2}{3} \\ -1 & \frac{1}{3} \end{bmatrix}$$

**Example.** *Find the inverse for the following matrix:* $\begin{bmatrix} 3 & -\frac{2}{3} \\ -1 & \frac{1}{3} \end{bmatrix}$

**Solution**

$$\left[\begin{array}{cccc} 3 & -\frac{2}{3} & \vdots & y_1 \\ -1 & \frac{1}{3} & \vdots & y_2 \end{array}\right] \div 3$$

$$\longrightarrow \left[\begin{array}{cccc} 1 & -\frac{2}{9} & \vdots & \frac{1}{3}y_1 \\ -1 & \frac{1}{3} & \vdots & y_2 \end{array}\right] +(I)$$

$$\longrightarrow \left[\begin{array}{cccc} 1 & -\frac{2}{9} & \vdots & \frac{1}{3}y_1 \\ 0 & \frac{1}{9} & \vdots & \frac{1}{3}y_1 + y_2 \end{array}\right] \times 9$$

$$\longrightarrow \left[\begin{array}{cccc} 1 & -\frac{2}{9} & \vdots & \frac{1}{3}y_1 \\ 0 & 1 & \vdots & 3y_1 + 9y_2 \end{array}\right] +\frac{2}{9}(II)$$

$$\longrightarrow \left[\begin{array}{cccc} 1 & 0 & \vdots & y_1 + 2y_2 \\ 0 & 1 & \vdots & 3y_1 + 9y_2 \end{array}\right]$$

$$\Rightarrow \begin{bmatrix} 3 & -\frac{2}{3} \\ -1 & \frac{1}{3} \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 2 \\ 3 & 9 \end{bmatrix}$$

**Example.** *Not all linear transformations are invertible. Consider the matrix* $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$

If $\vec{y} = \begin{bmatrix} 89 \\ 178 \end{bmatrix}$, to solve the system

$$\left| \begin{array}{rrr} x_1 & +2x_2 = & 89 \\ 2x_1 & +4x_2 = & 178 \end{array} \right|$$

We discover there are infinitely many solutions

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 89 - 2t \\ t \end{bmatrix}$$

We say that the coding matrix $A$ are *noninvertible*.

**Homework.** *Exercises 2.1: 4, 5, 7, 10, 12, 15*

## 2.2 Linear Transformation in Geometry

**Example. 1** *Consider a linear transformation system* $T(\vec{x}) = A\vec{x}$ *from* $R^n$ *to* $R^m$.

*a.* $T(\vec{v} + \vec{w}) = T(\vec{v}) + T(\vec{w})$

*In words, the transformation of the sum of two vectors equals the sum of the transformation.*

*b.* $T(k\vec{v}) = kT(\vec{v})$

*In words, the transformation of a scalar multiple of a vector is the scalar multiple of the transform.*

See Figure 1 (pp.50).

**Fact** A transformation $T$ from $R^n$ to $R^m$ is linear iff

a. $T(\vec{v}+\vec{w}) = T(\vec{v}) + T(\vec{w})$, for all $\vec{v}$, $\vec{w}$ in $R^n$, and

b. $T(k\vec{v}) = kT(\vec{v})$, for all $\vec{v}$ in $R^n$ and all scalars $k$.

**Proof**

Idea: To prove the inverse, we must show a matrix $A$ such that $T(\vec{x}) = A\vec{x}$. Consider a transformation $T$ from $R^n$ to $R^m$ that satisfy (a) and (b), find $A$.

**Example. 2** *Consider a linear transformation* $T$ *from* $R^2$ *to* $R^2$. *The vectors* $T\vec{e}_1$ *and* $T\vec{e}_2$ *are sketched in Figure 2. Sketch the* **image** *of the unit square under this transformation.*

*See Figure 2. (pp. 51)*

**Example. 3** *Consider a linear transformation* $T$ *from* $R^2$ *to* $R^2$ *such that* $T(\vec{v}_1) = \frac{1}{2}\vec{v}_1$ *and* $T(\vec{v}_2) = 2\vec{v}_2$, *for the vectors* $\vec{v}_1$ *and* $\vec{v}_2$ *in Figure 5. On the same axes, sketch* $T(\vec{x})$, *for the given vector* $\vec{x}$.

*See Figure 5. (pp. 52)*

**[Rotation]**

**Example. 4** *Let* $T$ *be the counterclockwise rotation through an angle* $\alpha$.

*a. Draw sketches to illustrate that* $T$ *is a linear transformation.*

*b. Find the matrix of* $T$.

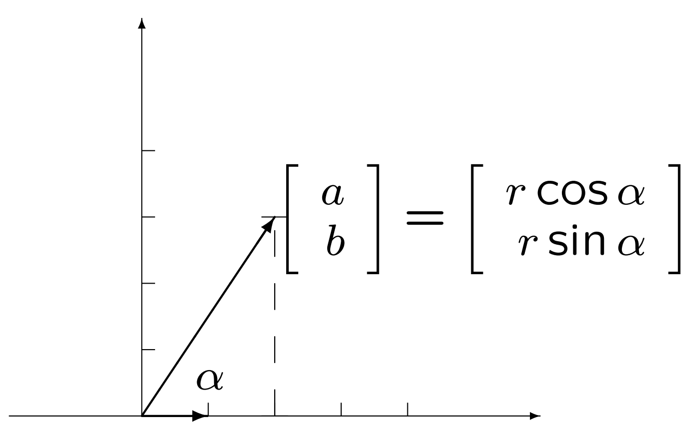
**Example. 5** *Give a geometric interpretation of the linear transformation.*

$$T(\vec{x}) = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \vec{x}$$

**Rotation-dilations** A matrix with this form

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

denotes a counterclockwise rotation through the anle $\alpha$ followed by a dilation by the factor $r$ where $\tan(\alpha) = \frac{b}{a}$ and $r = \sqrt{a^2 + b^2}$. Geometrically,



**[Shears]**

**Example. 6** *Consider the linear transformation*

$$T(\vec{x}) = \begin{bmatrix} 1 & \frac{1}{2} \\ 0 & 1 \end{bmatrix} \vec{x}$$

*To understand this transformation, sketch the image of the* **unit square***.*

**Solution** The transformation $T(\vec{x}) = \begin{bmatrix} 1 & \frac{1}{2} \\ 0 & 1 \end{bmatrix} \vec{x}$ is called a *shear* parallel to the $x_1$-axis.

**Definition. Shear** *Let $L$ be a line in $R^2$. A linear transformation $T$ from $R^2$ to $R^2$ is called a* **shear parallel** *to $L$ if*

*a. $T(\vec{v}) = \vec{v}$, for all vectors $\vec{v}$ on $L$, and*

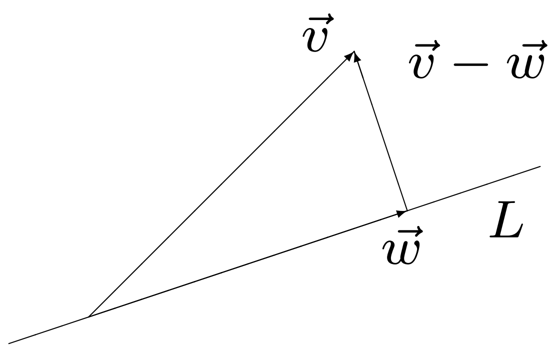
*b. $T(\vec{v}) - \vec{v}$ is parallel to $L$ for all vectors $\vec{x} \in R^2$.*

**Example. 7** *Consider two perpendicular vectors $\vec{u}$ and $\vec{w}$ in $R^2$. Show that the transformation*

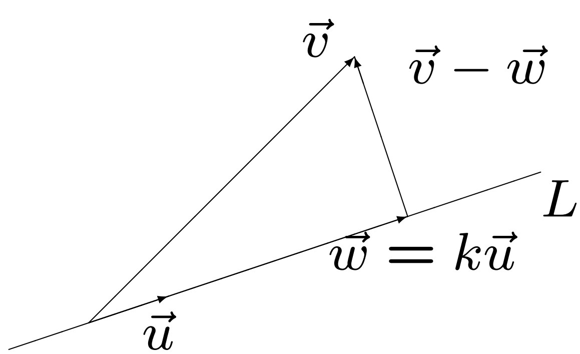
$$T(\vec{x}) = \vec{x} + (\vec{u} \cdot \bar{x})\vec{w}$$

*is a shear parallel to the line $L$ spanned by $\vec{w}$.*

Consider a line $L$ in $R^2$. For any vector $\vec{v}$ in $R^2$, there is a unique vector $\vec{w}$ on $L$ such that $\vec{v} - \vec{w}$ is perpendicular to $L$.




How can we generalize the idea of an orthogonal projection to lines in $R^n$?



**Definition. orthogonal projection** *Let $L$ be a line in $R^n$ consisting of all scalar multiples of some unit vector $\vec{u}$. For any vector $\vec{v}$ in $R^n$ there is a unique vector $\vec{w}$ on $L$ such that $\vec{v} - \vec{w}$ is perpendicular to $L$, namely, $\vec{w} = (\vec{u} \cdot \vec{v})\vec{u}$. This vector $\vec{w}$ is called the* **orthogonal projection** *of $\vec{v}$ onto $L$:*

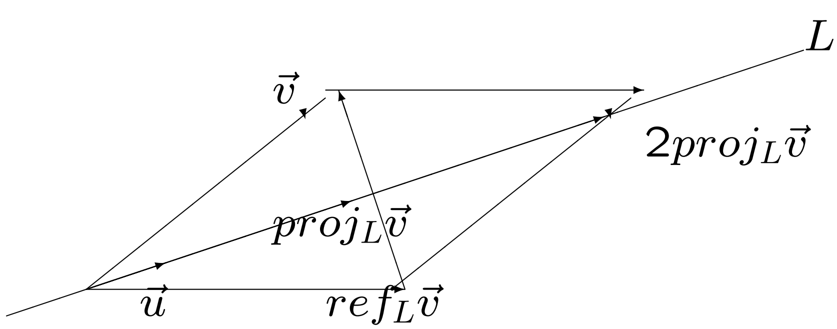
$$proj_L(\vec{v}) = (\vec{u} \cdot \vec{v})\vec{u}$$

*The transformation $proj_L$ from $R^n$ to $R^n$ is linear.*

**Definition.** *Let $L$ be a line in $R^n$, the vector $2(proj_L\vec{v}) - \vec{v}$ is called the* **reflection** *of $\vec{v}$ in $L$:*

$$ref_L(\vec{v}) = 2(proj_L\vec{v}) - \vec{v} = 2(\vec{u}\cdot\vec{v})\vec{u} - \vec{v}$$

*where $\vec{u}$ is a unit vector on $L$.*




**Homework.** *Exercise 2.2: 1, 9, 13, 17, 27*

## 2.3 The Inverse Of a Linear Transformation

**Definition.** *A function T from X to Y is called* **invertible** *if the equation T(x)=y has a* **unique solution x** *in X for each y in Y.*

Denote the inverse of $T$ as $T^{-1}$ from $Y$ to $X$, and write

$T^{-1}(y)$=(the unique $x$ in $X$ such that $T(x) = y$)

Note

$T^{-1}(T(x)) = x$, for all $x$ in $X$, and

$T(T^{-1}(y)) = y$, for all $y$ in $Y$.

If a function $T$ is invertible, then so is $T^{-1}$,

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

Consider the case of a *linear transformation* from $R^n$ to $R^m$ given by $\boxed{\vec{y} = A\vec{x}}$ where $A$ is an $m \times n$ matrix, the transformation is invertible if the linear system $\boxed{A\vec{x} = \vec{y}}$ has a unique solution.

1. **Case 1:** $m < n$ The system $A\vec{x} = \vec{y}$ has either no solutions or infinitely many solutions, for any $\vec{y}$ in $R^m$. Therefore $\vec{y} = A\vec{x}$ is noninvertible.

2. **Case 2:** $m = n$ The system $A\vec{x} = \vec{y}$ has a unique solution iff $rref(A) = I_n$, or equivalently, if $rank(A) = n$.

3. **Case 3:** $m > n$ The transformation $\vec{y} = A\vec{x}$ is noninvertible, because we can find a vector $\vec{y}$ in $R^m$ such that the system $A\vec{x} = \vec{y}$ is inconsistent.

**Definition. Invertible Matrix** *A matrix A is called invertible if the linear transformation $\vec{y} = A\vec{x}$ is invertible. The matrix of inverse transformation is denoted by $A^{-1}$. If the transformation $\vec{y} = A\vec{x}$ is invertible. its inverse is $\vec{x} = A^{-1}\vec{y}$.*

**Fact**

An $m \times n$ matrix A is invertible if and only if

1. $A$ is a square matrix (*i.e.,m=n*), and

2. $rref(A) = I_n$.

**Example.** *Is the matrix $A$ invertible?*

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

**Solution**

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \begin{matrix} \longrightarrow \\ -4(I) \\ -7(I) \end{matrix} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -12 \end{bmatrix} \begin{matrix} \longrightarrow \\ \div(-3) \\ \phantom{x} \end{matrix}$$

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & -6 & -12 \end{bmatrix} \begin{matrix} -2(II) \\ \longrightarrow \\ +6(II) \end{matrix} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$$

$A$ fails to be invertible, since $rref(A) \neq I_3$.

**Fact** Let $A$ be an $n \times n$ matrix.

1. Consider a vector $\vec{b}$ in $R^n$. If $A$ is invertible, then the system $A\vec{x} = \vec{b}$ has the unique solution $\vec{x} = A^{-1}\vec{b}$. If $A$ is noninvertible, then the system $A\vec{x} = \vec{b}$ has infinitely many solutions or none.

2. Consider the special case when $\vec{b} = \vec{0}$. The system $A\vec{x} = \vec{0}$. has $\vec{x} = \vec{0}$ as a solution. If $A$ is invertible, then this is the only solution. If $A$ is noninvertible, then there are infinitely many other solutions.

If a matrix $A$ is invertible, how can we find the inverse matrix $A^{-1}$?

Consider the matrix

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 8 & 2 \end{bmatrix}.$$

or, equivalently, the linear transformation

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} x_1 & + & x_2 & + & x_3 \\ 2x_1 & + & 3x_2 & + & 2x_3 \\ 3x_1 & + & 8x_2 & + & 2x_3 \end{bmatrix}.$$

To find the inverse transformation, we solve this system for input variables $x_1$, $x_2$, $x_3$:

$$\left| \begin{array}{rcrcrcl} x_1 & + & x_2 & + & x_3 & = & y_1 \\ 2x_1 & + & 3x_2 & + & 2x_3 & = & \qquad y_2 \\ 3x_1 & + & 8x_2 & + & 2x_3 & = & \qquad\qquad y_3 \end{array} \right| \begin{array}{r} \longrightarrow \\ -2(I) \\ -3(I) \end{array}$$

$$\left| \begin{array}{rcrcrcrcrcr} x_1 & + & x_2 & + & x_3 & = & y_1 & & & & \\ & & x_2 & & & = & -2y_1 & + & y_2 & & \\ & & 5x_2 & - & 3x_3 & = & -3y_1 & & & + & y_3 \end{array} \right| \begin{array}{r} -(II) \\ \longrightarrow \\ -5(II) \end{array}$$

$$\left| \begin{array}{rrcrcrcrcr} x_1 & & + & x_3 & = & 3y_1 & - & y_2 & & \\ & x_2 & & & = & -2y_1 & + & y_2 & & \\ & & - & x_3 & = & 7y_1 & - & 5y_2 & + & y_3 \end{array} \right| \begin{array}{r} \longrightarrow \\ \\ \div(-1) \end{array}$$

$$\left| \begin{array}{rrcrcrcrcr} x_1 & & + & x_3 & = & 3y_1 & - & y_2 & & \\ & x_2 & & & = & -2y_1 & + & y_2 & & \\ & & & x_3 & = & -7y_1 & + & 5y_2 & - & y_3 \end{array} \right| \begin{array}{r} -(III) \\ \longrightarrow \\ \\ \end{array}$$

$$\left| \begin{array}{rrrcrcrcr} x_1 & & & = & 10y_1 & - & 6y_2 & + & y_3 \\ & x_2 & & = & -2y_1 & + & y_2 & & \\ & & x_3 & = & -7y_1 & + & 5y_2 & - & y_3 \end{array} \right| .$$

We have found the inverse transformation; its matrix is

$$B = A^{-1} = \begin{bmatrix} 10 & -6 & 1 \\ -2 & 1 & 0 \\ -7 & 5 & -1 \end{bmatrix}.$$

We can write the preceding computations in matrix form:

$$\left[\begin{array}{ccccccc} 1 & 1 & 1 & : & 1 & 0 & 0 \\ 2 & 3 & 2 & : & 0 & 1 & 0 \\ 3 & 8 & 2 & : & 0 & 0 & 1 \end{array}\right] \begin{array}{r} \longrightarrow \\ -2(I) \\ -3(I) \end{array}$$

$$\left[\begin{array}{ccccccc} 1 & 1 & 1 & : & 1 & 0 & 0 \\ 0 & 1 & 0 & : & -2 & 1 & 0 \\ 0 & 5 & -1 & : & -3 & 0 & 1 \end{array}\right] \begin{array}{r} -(II) \\ \longrightarrow \\ -5(II) \end{array}$$

$$\left[\begin{array}{ccccccc} 1 & 0 & 1 & : & 3 & -1 & 0 \\ 0 & 1 & 0 & : & -2 & 1 & 0 \\ 0 & 0 & -1 & : & 7 & -5 & 1 \end{array}\right] \begin{array}{r} \longrightarrow \\ \\ \div(-1) \end{array}$$

$$\left[\begin{array}{ccccccc} 1 & 0 & 1 & : & 3 & -1 & 0 \\ 0 & 1 & 0 & : & -2 & 1 & 0 \\ 0 & 0 & 1 & : & -7 & 5 & -1 \end{array}\right] \begin{array}{r} -(III) \\ \longrightarrow \\ \\ \end{array}$$

$$\left[\begin{array}{ccccccc} 1 & 0 & 0 & : & 10 & -6 & 1 \\ 0 & 1 & 0 & : & -2 & 1 & 0 \\ 0 & 0 & 1 & : & -7 & 5 & -1 \end{array}\right].$$

This process can be described succinctly as follows:

## Find the inverse of a matrix

To find the inverse of an $n \times n$ matrix $A$, from the $n \times (2n)$ matrix $\left[ \begin{array}{ccc} A & \vdots & I_n \end{array} \right]$ and compute rref $\left[ \begin{array}{ccc} A & \vdots & I_n \end{array} \right]$.

- If rref $[A \vdots I_n]$ is of the form $[I_n \vdots B]$, then $A$ is invertible, and $A^{-1} = B$.

- If rref $[A \vdots I_n]$ is of another form (*i.e., its left half fails to be* $I_n$), then $A$ is not invertible. (*Note that the left half of rref* $[A \vdots I_n]$ *is rref(A).*)

The inverse of a 2×2 matrix is particularly easy to find.

**Inverse and determinant of a $2 \times 2$ matrix**

1. The $2 \times 2$ matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

   is invertible if (and only if) $ad - bc \neq 0$. Quantity $\boxed{ad - bc}$ is called the determinant of $A$, written det($A$):

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

2. If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is invertible, then

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{\mathrm{ad-bc}} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \frac{1}{\mathrm{det(A)}} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

   Compare this with Exercise 2.1.13.

**Homework.** *Exercise 2.3 21–27, 41*

# 2.4 MATRIX PRODUCTS

The *composite* of two functions: $y = sin(x)$ and $z = cos(y)$ is $z = cos(sin(x))$.

Consider two transformation systems:

$$\vec{y} = A\vec{x}, \text{ with } A = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}$$

$$\vec{z} = B\vec{y}, \text{ with } B = \begin{bmatrix} 6 & 7 \\ 8 & 9 \end{bmatrix}$$

The composite of the two transformation systems is

$$\vec{z} = B(A\vec{x})$$

Question: Is $\vec{z} = T(\vec{x})$ linear? If so, what's the matrix?

(a) Find the matrix for the composite:

$$\begin{aligned} z_1 &= 6y_1 + 7y_2 \\ z_2 &= 8y_1 + 9y_2 \end{aligned} \quad \textit{and} \quad \begin{aligned} y_1 &= \phantom{3}x_1 + 2x_2 \\ y_2 &= 3x_1 + 5x_2 \end{aligned}$$

$$\begin{aligned} &z_1 = 6(x_1 + 2x_2) + 7(3x_1 + 5x_2) \\ &= (6 \cdot 1 + 7 \cdot 3)x_1 + (6 \cdot 2 + 7 \cdot 5)x_2 \\ &= 27x_1 + 47x_2 \end{aligned}$$

$$\begin{aligned} &z_2 = 8(x_1 + 2x_2) + 9(3x_1 + 5x_2) \\ &= (8 \cdot 1 + 9 \cdot 3)x_1 + (8 \cdot 2 + 9 \cdot 5)x_2 \\ &= 35x_1 + 61x_2 \end{aligned}$$

This shows the composite is linear with matrix

$$\begin{bmatrix} 6 \cdot 1 + 7 \cdot 3 & 6 \cdot 2 + 7 \cdot 5 \\ 8 \cdot 1 + 9 \cdot 3 & 8 \cdot 2 + 9 \cdot 5 \end{bmatrix} = \begin{bmatrix} 27 & 47 \\ 35 & 61 \end{bmatrix}$$

(b)Use Fact to show the transformation $T(\vec{x}) = B(A\vec{x})$ is linear:

$T(\vec{v}+\vec{w}) = B(A(\vec{v}+\vec{w})) = B(A\vec{v}+A\vec{w}) = B(A\vec{v}) + B(A\vec{w}) = T(\vec{v}) + T(\vec{w})$

$T(k\vec{v}) = B(A(k\vec{v})) = B(k(A\vec{v})) = k(B(A\vec{v})) = kT(\vec{v})$

Once we know that $T$ is linear, we can find its matrix by computing the vectors: $T(\vec{e}_1)$ and $T(\vec{e}_2)$:

$T(\vec{e}_1) = B(A(\vec{e}_1))=B$(first column of $A$) $= \begin{bmatrix} 6 & 7 \\ 8 & 9 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 27 \\ 35 \end{bmatrix}$

$T(\vec{e}_2) = B(A(\vec{e}_1))=B$(second column of $A$) $= \begin{bmatrix} 6 & 7 \\ 8 & 9 \end{bmatrix} \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 47 \\ 61 \end{bmatrix}$

The matrix of $T(\vec{x}) = B(A\vec{x}) = BA(\vec{x})$:

$$= \begin{bmatrix} | & | \\ T(\vec{e}_1) & T(\vec{e}_2) \\ | & | \end{bmatrix} = \begin{bmatrix} 27 & 47 \\ 35 & 61 \end{bmatrix}$$

**Definition. Matrix multiplication**

1. Let $B$ be an $m \times n$ matrix and $A$ a $q \times p$ matrix. The product $BA$ is defined if (and only if) $n = q$.

2. If $B$ is an $m \times n$ matrix and $A$ an $n \times p$ matrix, then the product $BA$ is defined as the matrix of the linear transformation $T(\vec{x}) = B(A\vec{x})$. This means that $T(\vec{x}) = B(A\vec{x}) = (BA)\vec{x}$, for all $\vec{x}$ in $R^p$. The product $BA$ is an $m \times p$ matrix.

Let $B$ be an $m \times n$ matrix and A an $n \times p$ matrix. Let's think about the columns of the matrix $BA$:

($i$th columns of $BA$)
$= (BA)\vec{e}_i$
$= B(A\vec{e}_i)$
$= B$($i$th column of $A$).

If we denote the columns of A by $\vec{v}_1, \vec{v}_2, ..., \vec{v}_p$,we can write

$$BA = B \underbrace{\begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_p \\ | & | & & | \end{bmatrix}}_{A} =$$

$$\begin{bmatrix} | & | & & | \\ B\vec{v}_1 & B\vec{v}_2 & \dots & B\vec{v}_p \\ | & | & & | \end{bmatrix}.$$

## The matrix product, column by column

Let $B$ be an $m \times n$ matrix and $A$ an $n \times p$ matrix with columns $\vec{v}_1, \vec{v}_2, ..., \vec{v}_p$. Then, the product $BA$ is

$$BA = B \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \ldots & \vec{v}_p \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ B\vec{v}_1 & B\vec{v}_2 & \ldots & B\vec{v}_p \\ | & | & & | \end{bmatrix}.$$

To find $BA$, we can multiply $B$ with the columns of $A$ and combine the resulting vectors.

**Fact** Matrix multiplication is noncommutative: $AB \neq BA$, in general. However, at times it does happen that $AB = BA$ ; then, we say that the matrices $A$ and $B$ commute.

## The matrix product, entry by entry

Let $B$ be an $m \times n$ matrix and $A$ an $n \times p$ matrix. The $ij$th entry of $BA$ is the dot product of the $i$th row of $B$ and the $j$th column of $A$.

$$\begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{i1} & b_{i2} & \dots & b_{in} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nj} & \dots & a_{np} \end{bmatrix}$$

is the $m \times p$ matrix whose $ij$th entry is

$$b_{i1}a_{1j} + b_{i2}a_{2j} + \ldots + b_{in}a_{nj} = \sum_{k=1}^{n} b_{ik}a_{kj}.$$

**Example.** $\begin{bmatrix} 6 & 7 \\ 8 & 9 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix} =$

$$\begin{bmatrix} 6 \cdot 1 + 7 \cdot 3 & 6 \cdot 2 + 7 \cdot 5 \\ 8 \cdot 1 + 9 \cdot 3 & 8 \cdot 2 + 9 \cdot 5 \end{bmatrix} = \begin{bmatrix} 27 & 47 \\ 35 & 61 \end{bmatrix}.$$

*We have done these computations before.(where?)*

## Matrix Algebra

**Fact** For an invertible $n \times n$ matrix A.

$$AA^{-1} = I_n \text{ and } A^{-1}A = I_n.$$

**Fact** For an $m \times n$ matrix A.

$$AI_n = I_m A = A.$$

**Fact** Matrix multiplication is associative

$$(AB)C = A(BC).$$

We can write simply $ABC$ for the product $(AB)C = A(BC)$.

**Proof** (a) $(AB)C = (AB)[\ \vec{v}_1 \ \ \vec{v}_2 \ \ \cdots \ \ \vec{v}_q\ ]$

$$= [\ (AB)\vec{v}_1 \ \ (AB)\vec{v}_2 \ \ \cdots \ \ (AB)\vec{v}_q\ ]$$

and

$$A(BC) = A[\ B\vec{v}_1 \ \ B\vec{v}_2 \ \ \cdots \ \ B\vec{v}_q\ ]$$

$$= [\ A(B\vec{v}_1) \ \ A(B\vec{v}_2) \ \ \cdots \ \ A(B\vec{v}_q)\ ]$$

Since $(AB)\vec{v}_i = A(B\vec{v}_i)$, by definition of the matrix product, we find that $(AB)C = A(BC)$.

**Proof** (b) Consider two linear transformations

$$T(\vec{x}) = ((AB)C)\vec{x}$$

and

$$L(\vec{x}) = (A(BC))\vec{x}$$

are identical because,

$$T(\vec{x}) = ((AB)C)\vec{x} = (AB)(C\vec{x}) = A(B(C\vec{x}))$$

and

$$L(\vec{x}) = (A(BC))\vec{x} = A((BC)\vec{x}) = A(B(C\vec{x}))$$

If $A$ and $B$ are invertible $n \times n$ matrices, is $BA$ invertible?

$$\vec{y} = BA\vec{x}$$

multiply both sides by $B^{-1}$

$$B^{-1}\vec{y} = B^{-1}BA\vec{x} = I_n A\vec{x} = A\vec{x}$$

next, multiply both sides by $A^{-1}$

$$A^{-1}B^{-1}\vec{y} = A^{-1}A\vec{x} = I_n\vec{x} = \vec{x}$$

This computation shows that the linear transformation is invertible since

$$\vec{x} = A^{-1}B^{-1}\vec{y}$$

**Fact** The inverse of a product of matrices

If $A$ and $B$ are invertible $n \times n$ matrices, then $BA$ is invertible as well, and

$$(BA)^{-1} = A^{-1}B^{-1}.$$

Pay attention to the order of the matrices.

**Proof** Verify it by yourself.

**Fact** Let $A$ and $B$ be two $n \times n$ matrices such that

$$BA = I_n.$$

Then,

a. $A$ and $B$ are both invertible.
b. $A^{-1} = B$ and $B^{-1} = A$, and
c. $AB = I_n$.

**Proof** (a) To demonstrate $A$ is invertible it suffices to show that the linear system $A\vec{x} = \vec{0}$ has only the solution $\vec{x} = \vec{0}$.

$$BA\vec{x} = B\vec{0} = \vec{0}$$

(b) $B = A^{-1}$ since

$$(BA)A^{-1} = (I_n)A^{-1} = A^{-1}$$

and

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

(c) $AB = AA^{-1} = I_n$

**Example.** *B*= $\frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$ *is the inverse of* $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

*it suffices to verify that* $BA = I_2$*:*

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

$$= \frac{1}{ad-bc}\begin{bmatrix} ad-bc & bd-bd \\ ac-ac & ad-bc \end{bmatrix} = I_2.$$

**Example.**

Suppose $A$, $B$ and $C$ are three $n \times n$ matrices and $ABC = I_n$. Show that $B$ is invertible, and express $B^{-1}$ in term of $A$ and $C$.

**Solution**

Write $ABC = (AB)C = I_n$. We have $C(AB) = I_n$. Since matrix multiplication is associative, we can write $(CA)B = I_n$. We conclude that $B$ is invertible, and $B^{-1} = CA$.

**Distributive property for matrices**

**Fact** If $A$, $B$ are $n \times n$, and $C$, $D$ are $n \times p$ matrices, then

$$A(C + D) = AC + AD$$

and

$$(A + B)C = AC + BC.$$

**Fact** If $A$ is an $m \times n$ matrix, $B$ an $n \times p$ matrix, and $k$ a scalar, then

$$(kA)B = A(kB) = k(AB).$$

## Partitioned Matrices

It is sometimes useful to break a large matrix down into smaller submatrices by slicing it up with horizontal or vertical lines that go all the way through the matrix.

For example, we can think of the $4 \times 4$ matrix

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

as a $2 \times 2$ matrix whose "entries" are four $2 \times 2$ matrices:

$$A = \left[\begin{array}{cc|cc} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ \hline 9 & 8 & 7 & 6 \\ 5 & 4 & 3 & 2 \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

with $A_{11} = \begin{bmatrix} 1 & 2 \\ 5 & 7 \end{bmatrix}$, $A_{12} = \begin{bmatrix} 3 & 4 \\ 7 & 8 \end{bmatrix}$, etc.

The submatrices in such a partition need not be of equal size; for example, we could have

$$B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \left[\begin{array}{cc|c} 1 & 2 & 3 \\ 4 & 5 & 6 \\ \hline 7 & 8 & 9 \end{array}\right] = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}.$$

A useful property of partitioned matrices is the following:

**Multiplying partitioned matrices** Partitioned matrices can be multiplied as though the submatrices were scalars:

$AB =$

$$\begin{bmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{i1} & A_{i2} & \dots & A_{in} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} & \dots & B_{1j} & \dots & B_{1p} \\ B_{21} & B_{22} & \dots & B_{2j} & \dots & B_{2p} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ B_{n1} & B_{n2} & \dots & B_{nj} & \dots & B_{np} \end{bmatrix}$$

is the partitioned matrix whose ijth "entry" is the matrix

$$A_{i1}B_{1j} + A_{i2}B_{2j} + \ldots + A_{in}B_{nj} = \sum_{k=1}^{n} A_{ik}B_{kj},$$

provided that all the products $A_{ik}B_{kj}$ are defined.

**Example.**

$$\left[\begin{array}{cc|c} 0 & 1 & -1 \\ 1 & 0 & 1 \end{array}\right] \left[\begin{array}{cc|c} 1 & 2 & 3 \\ 4 & 5 & 6 \\ \hline 7 & 8 & 9 \end{array}\right]$$

$$= \left[\begin{array}{c|c} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix} + \begin{bmatrix} -1 \\ 1 \end{bmatrix}\begin{bmatrix} 7 & 8 \end{bmatrix} & \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 3 \\ 6 \end{bmatrix} + \begin{bmatrix} -1 \\ 1 \end{bmatrix}[9] \end{array}\right]$$

$$= \left[\begin{array}{cc|c} -3 & -3 & -3 \\ 8 & 10 & 12 \end{array}\right].$$

Compute this product without using a partition, and see whether you find the same result.

**Example.**

$$A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix},$$

where $A_{11}$ is an $n \times n$ matrix, $A_{22}$ is an $m \times m$ matrix, and $A_{12}$ is an $n \times m$ matrix.

a. For which choices of $A_{11}$, $A_{12}$, and $A_{22}$ is $A$ invertible ?

b. If $A$ is invertible, what is $A^{-1}$ (in terms of $A_{11}$, $A_{12}$, $A_{22}$)?

**Solution**

We are looking for an $(n+m)\times(n+m)$ matrix $B$ such that

$$BA = I_{n+m} = \begin{bmatrix} I_n & 0 \\ 0 & I_m \end{bmatrix}.$$

Let us partition $B$ in the same way as $A$:

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix},$$

where $B_{11}$ is $n\times n$, $B_{22}$ is $m\times m$, etc. The fact that $B$ is the inverse of $A$ means that

$$\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ 0 & B_{22} \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ 0 & I_m \end{bmatrix},$$

or using

$$\left| \begin{array}{rrcr} & B_{11}A_{11} & = & I_n \\ B_{11}A_{12} & +B_{12}A_{22} & = & 0 \\ & B_{21}A_{11} & = & 0 \\ B_{21}A_{12} & +B_{22}A_{22} & = & I_m \end{array} \right|.$$

We have to solve this system for the submatrices $B_{ij}$.

1. By Equation 1, $A_{11}$ must be invertible, and $B_{11} = A_{11}^{-1}$.

2. By Equation 3, $B_{21} = 0$ (Multiply by $A_{11}^{-1}$ form the right)

3. Equation 4 now simplifies to $B_{22}A_{22} = I_m$. Therefore, $A_{22}$ must be invertible, and $B_{22} = A_{22}^{-1}$.

4. Lastly, Solve for $B_{12}$ by Equation 2

$$A_{11}^{-1}A_{12} + B_{12}A_{22} = 0$$

$$\Rightarrow B_{12}A_{22} = -A_{11}^{-1}A_{12}$$

$$\Rightarrow B_{12} = -A_{11}^{-1}A_{12}A_{22}^{-1}$$

So

a. $A$ is invertible if (and only if) both $A_{11}$ and $A_{22}$ are invertible (no condition is imposed on $A_{12}$).

b. If $A$ is invertible, then its inverse is

$$A^{-1} = \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ 0 & A_{22}^{-1} \end{bmatrix}.$$

Verify this result for the following example:

**Example.** *5*

$$\left[\begin{array}{cc|ccc} 1 & 1 & 1 & 2 & 3 \\ 1 & 2 & 4 & 5 & 6 \\ \hline 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right]^{-1} = \left[\begin{array}{cc|ccc} 2 & -1 & 2 & 1 & 0 \\ -1 & 1 & -3 & -3 & -3 \\ \hline 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right].$$

**Homework.**

Exercise 2.4: 5, 13, 17, 23, 27, 35