Linear Transformations

Ryan C. Daileda

Trinity University

Linear Algebra

Matrix Transformations

Let A be an $m \times n$ matrix.

Given a vector $\mathbf{x} \in \mathbb{R}^n$, the matrix-vector product $A\mathbf{x}$ is a vector in $\mathbb{R}^m$.

The rule $T(\mathbf{x}) = A\mathbf{x}$ therefore defines a *matrix transformation* $T : \mathbb{R}^n \to \mathbb{R}^m$.

Recall that $\mathbb{R}^n$ is the *domain* of T, $\mathbb{R}^m$ is the *codomain* of T, and the *image* (or *range*) of T is

$$\text{im}\, T = \{T(\mathbf{x}) \,|\, \mathbf{x} \in \mathbb{R}^n\} = \{A\mathbf{x} \,|\, \mathbf{x} \in \mathbb{R}^n\} = \text{Col}\, A.$$

Example 1

Let

$$A = \begin{pmatrix} 1 & -5 & -7 \\ -3 & 7 & 5 \end{pmatrix}, \quad \mathbf{u} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} -2 \\ -2 \end{pmatrix}.$$

Define $T : \mathbb{R}^3 \to \mathbb{R}^2$ by $T(\mathbf{x}) = A\mathbf{x}$.

(a) Compute $T(\mathbf{u})$.

(b) Find a vector $\mathbf{x} \in \mathbb{R}^3$ so that $T(\mathbf{x}) = \mathbf{b}$ (a *preimage* of $\mathbf{b}$). Is $\mathbf{x}$ unique?

(c) Determine $\operatorname{im} T$.

Solution. (a) We have

$$T(\mathbf{u}) = A\mathbf{u} = 1\begin{pmatrix} 1 \\ -3 \end{pmatrix} + 2\begin{pmatrix} -5 \\ 7 \end{pmatrix} + 3\begin{pmatrix} -7 \\ 5 \end{pmatrix} = \begin{pmatrix} -30 \\ 26 \end{pmatrix}.$$

(b) To solve $T(\mathbf{x}) = A\mathbf{x} = \mathbf{b}$ we row reduce the augmented matrix:

$$(A \quad \mathbf{b}) = \begin{pmatrix} 1 & -5 & -7 & -2 \\ -3 & 7 & 5 & -2 \end{pmatrix} \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 0 & 3 & 3 \\ 0 & 1 & 2 & 1 \end{pmatrix}.$$

This tells us that

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 - 3x_3 \\ 1 - 2x_3 \\ x_3 \end{pmatrix}.$$

Any choice of x_3 yields a preimage of $\mathbf{b}$, for instance

$$\mathbf{x} = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}.$$

Because x_3 is free, $\mathbf{x}$ is *not unique*.

(c) Because

$$A \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \end{pmatrix},$$

we see that A has two pivot columns.

Thus $\dim \operatorname{Col} A = 2$. But $\operatorname{Col} A \leq \mathbb{R}^2$, so this means

$$\operatorname{im} T = \operatorname{Col} A = \mathbb{R}^2.$$

$\square$

In general, questions/statements about the $m \times n$ matrix equation $A\mathbf{x} = \mathbf{b}$ can be reformulated in terms of the matrix transformation $T(\mathbf{x}) = A\mathbf{x}$.

$T(\mathbf{x}) = A\mathbf{x}$	$A\mathbf{x} = \mathbf{b}$
Is $\mathbf{b}$ in im T?	Does $A\mathbf{x} = \mathbf{b}$ have a solution?
Is T onto?	Does Col $A = \mathbb{R}^m$?
Is T one-to-one?	Are the solutions to $A\mathbf{x} = \mathbf{b}$ unique?

As we have seen, we can answer all of the questions on the right by row reducing either A or the augmented matrix $\begin{pmatrix} A & \mathbf{b} \end{pmatrix}$.

Linear Transformations

If A is an $m \times n$ matrix, and $T(\mathbf{x}) = A\mathbf{x}$, the properties of matrix-vector multiplication show that for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $c \in \mathbb{R}$ we have:

$$T(\mathbf{x}+\mathbf{y}) = A(\mathbf{x}+\mathbf{y}) = A\mathbf{x} + A\mathbf{y} = T(\mathbf{x}) + T(\mathbf{y}),$$

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

Functions with these two properties are particularly important in linear algebra.

Definition

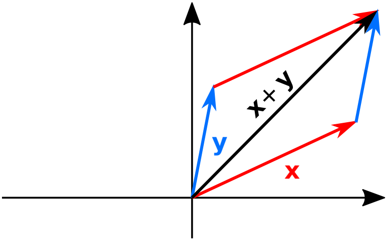
A function $T : \mathbb{R}^n \to \mathbb{R}^m$ is called a *linear transformation* iff:

- $T(\mathbf{x}+\mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,
- $T(c\mathbf{x}) = cT(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$ and $c \in \mathbb{R}$.

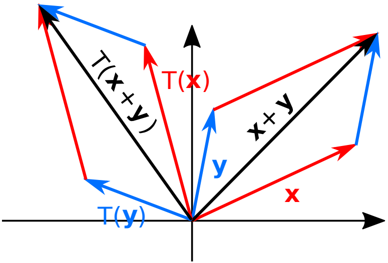
Example

Fix an angle θ and let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be (counterclockwise) rotation about the origin by θ.

Given $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$, the tip-to-tail rule shows that $\mathbf{x} + \mathbf{y}$ is the diagonal of the parallelogram determined by $\mathbf{x}$ and $\mathbf{y}$.



If we rotate this parallelogram about the origin, we find that $T(\mathbf{x}+\mathbf{y})$ is the diagonal of the parallelogram determined by $T(\mathbf{x})$ and $T(\mathbf{y})$.



The tip-to-tail rule then shows that $T(\mathbf{x}+\mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$.

If $c \in \mathbb{R}$, then $\mathbf{x}$ and $c\mathbf{x}$ have the same (or opposite) direction.

Therefore if we rotate $\mathbf{x}$ and $c\mathbf{x}$ by the same amount, they will still have the same direction.

So $T(\mathbf{x})$ and $T(c\mathbf{x})$ have the same direction.

Since rotation doesn't change lengths,

$$|T(c\mathbf{x})| = |c\mathbf{x}| = |c| \cdot |\mathbf{x}| = |c| \cdot |T(\mathbf{x})|.$$

We conclude that $T(c\mathbf{x})$ has the same direction as $T(\mathbf{x})$ and length $|c| \cdot |T(\mathbf{x})|$.

This is precisely the geometric description of $cT(\mathbf{x})$, so that

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

We conclude that:

Theorem 1

Rotation of $\mathbb{R}^2$ about the origin by a fixed angle θ is a linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$.

Now fix a scalar a and define $T : \mathbb{R}^n \to \mathbb{R}^n$ by

$$T(\mathbf{x}) = a\mathbf{x},$$

which we call *scaling by a*.

Based on our experience with $\mathbb{R}^2$ and $\mathbb{R}^3$, we call T *dilation* if $a > 1$ and a *contraction* if $0 < a < 1$.

If $a = 1$, we call T the *identity transformation*.

If $a = -1$ we call T *reflection through the origin*.

It is not hard to show that every scaling map is linear.

Indeed, if $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $c \in \mathbb{R}$, then

$$T(\mathbf{x}+\mathbf{y}) = a(\mathbf{x}+\mathbf{y}) = a\mathbf{x} + a\mathbf{y} = T(\mathbf{x}) + T(\mathbf{y}),$$
$$T(c\mathbf{x}) = a(c\mathbf{x}) = (ac)\mathbf{x} = (ca)\mathbf{x} = c(a\mathbf{x}) = cT(\mathbf{x}).$$

Let's record this fact.

Theorem 2

For any $a \in \mathbb{R}$, the scaling map $T : \mathbb{R}^n \to \mathbb{R}^n$ given by $T(\mathbf{x}) = a\mathbf{x}$ is a linear transformation.

Based on our earlier work with coordinates, we also have:

Theorem 3

Let $\mathcal{B}$ be a basis for $\mathbb{R}^n$. Then the coordinate map $[\cdot]_{\mathcal{B}} : \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation.

Let's look at a particular instance.

Example 2

If

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \end{pmatrix} \right\},$$

find a formula for $[\mathbf{x}]_{\mathcal{B}}$.

Solution. Recall that the $\mathcal{B}$-coordinate vector of $\mathbf{x} \in \mathbb{R}^2$ satisfies

$$\begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} [\mathbf{x}]_{\mathcal{B}} = \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

So to find $[\mathbf{x}]_{\mathcal{B}}$ we row reduce the augmented matrix:

$$\begin{pmatrix} 1 & 3 & x_1 \\ 2 & 4 & x_2 \end{pmatrix} \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 0 & -2x_1 + \frac{3}{2}x_2 \\ 0 & 1 & x_1 - \frac{1}{2}x_2 \end{pmatrix}.$$

This shows that $\mathcal{B}$ is indeed a basis for $\mathbb{R}^2$ (why?) and that

$$\left[\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right]_{\mathcal{B}} = \begin{pmatrix} -2x_1 + \frac{3}{2}x_2 \\ x_1 - \frac{1}{2}x_2 \end{pmatrix}.$$

$\square$

Non-Example

We don't have to try to hard to produce *non-linear* maps.
Fix a vector $\mathbf{v} \in \mathbb{R}^n$ and define the *translation map* $T : \mathbb{R}^n \to \mathbb{R}^n$ by

$$T(\mathbf{x}) = \mathbf{x} + \mathbf{v}.$$

Notice that

$$T(\mathbf{x} + \mathbf{y}) = (\mathbf{x} + \mathbf{y}) + \mathbf{v} = \mathbf{x} + \mathbf{v} + \mathbf{y} = T(\mathbf{x}) + \mathbf{y}.$$

This will equal $T(\mathbf{x}) + T(\mathbf{y})$ iff

$$\mathbf{y} = T(\mathbf{y}) = \mathbf{y} + \mathbf{v} \iff \mathbf{v} = \mathbf{0}.$$

That is, translation by a *nonzero* vector is *non-linear*.

Properties of Linear Transformations

Suppose $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation.

Notice that

$$T(\mathbf{0}) = T(\mathbf{0} + \mathbf{0}) = T(\mathbf{0}) + T(\mathbf{0}) \Rightarrow T(\mathbf{0}) = T(\mathbf{0}) - T(\mathbf{0}) = \mathbf{0}.$$

Also, for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and any $c, d \in \mathbb{R}$ we have

$$T(c\mathbf{x} + d\mathbf{y}) = T(c\mathbf{x}) + T(d\mathbf{y}) = cT(\mathbf{x}) + dT(\mathbf{y}).$$

More generally, one can use *mathematical induction* to show that the analogous result holds for linear combinations of arbitrary length.

Specifically, we have:

Theorem 4

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ *be a linear transformation. Then*

$$T(\mathbf{0}) = \mathbf{0},$$

and for any $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k \in \mathbb{R}^n$ *and any scalars* $c_1, c_2, \ldots, c_k \in \mathbb{R}$*:*

$$T(c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_k\mathbf{x}_k) = c_1 T(\mathbf{x}_1) + c_2 T(\mathbf{x}_2) + \cdots + c_k T(\mathbf{x}_k).$$

The second equation in Theorem 3 is sometimes called the *principle of superposition*.

We say that a linear transformation *respects* (or *preserves*) linear combinations.

Example

Example 3

Suppose $T : \mathbb{R}^2 \to \mathbb{R}^3$ is a linear transformation and we know that

$$T\begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix}-1\\0\\1\end{pmatrix} \quad \text{and} \quad T\begin{pmatrix}3\\4\end{pmatrix} = \begin{pmatrix}0\\-2\\3\end{pmatrix}.$$

Find a formula for $T(\mathbf{x})$.

Solution. We have seen that

$$\mathcal{B} = \left\{ \begin{pmatrix}1\\2\end{pmatrix}, \begin{pmatrix}3\\4\end{pmatrix} \right\}$$

is a basis for $\mathbb{R}^2$.

Let

$$A = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix},$$

so that $A[\mathbf{x}]_{\mathcal{B}} = \mathbf{x}$.

According to our work above

$$[\mathbf{x}]_{\mathcal{B}} = \begin{pmatrix} -2x_1 + \frac{3}{2}x_2 \\ x_1 - \frac{1}{2}x_2 \end{pmatrix},$$

so that

$$T(\mathbf{x}) = T(A[\mathbf{x}]_{\mathcal{B}}) = T\left(\left(-2x_1 + \frac{3}{2}x_2\right)\begin{pmatrix} 1 \\ 2 \end{pmatrix} + \left(x_1 - \frac{1}{2}x_2\right)\begin{pmatrix} 3 \\ 4 \end{pmatrix}\right)$$

$$= \left(-2x_1 + \frac{3}{2}x_2\right) T\begin{pmatrix} 1 \\ 2 \end{pmatrix} + \left(x_1 - \frac{1}{2}x_2\right) T\begin{pmatrix} 3 \\ 4 \end{pmatrix}$$

$$= \left(-2x_1 + \frac{3}{2}x_2\right)\begin{pmatrix}-1\\0\\1\end{pmatrix} + \left(x_1 - \frac{1}{2}x_2\right)\begin{pmatrix}0\\-2\\3\end{pmatrix}$$

$$= \begin{pmatrix}2x_1 - \frac{3}{2}x_2\\-2x_1 + x_2\\-2x_1 + \frac{3}{2}x_2 + 3\left(x_1 - \frac{1}{2}x_2\right)\end{pmatrix} = \begin{pmatrix}2x_1 - \frac{3}{2}x_2\\-2x_1 + x_2\\x_1\end{pmatrix}.$$

Thus

$$\boxed{T\begin{pmatrix}x_1\\x_2\end{pmatrix} = \begin{pmatrix}2x_1 - \frac{3}{2}x_2\\-2x_1 + x_2\\x_1\end{pmatrix}}.$$

□

The Matrix of a Linear Transformation

More generally, suppose $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a basis for $\mathbb{R}^n$ and $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation.

For $\mathbf{x} \in \mathbb{R}^n$ write

$$[\mathbf{x}]_{\mathcal{B}} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}.$$

Then

$$\begin{aligned} T(\mathbf{x}) &= T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n) \\ &= c_1 T(\mathbf{v}_1) + c_2 T(\mathbf{v}_2) + \cdots + c_n T(\mathbf{v}_n) \\ &= \begin{pmatrix} T(\mathbf{v}_1) & T(\mathbf{v}_2) & \cdots & T(\mathbf{v}_n) \end{pmatrix} [\mathbf{x}]_{\mathcal{B}}. \end{aligned}$$

This shows that if we know the values of T on the basis vectors in $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, then we can compute $T(\mathbf{x})$ by multiplying the coordinate vector $[\mathbf{x}]_{\mathcal{B}}$ by the matrix

$$\begin{pmatrix} T(\mathbf{v}_1) & T(\mathbf{v}_2) & \cdots & T(\mathbf{v}_n) \end{pmatrix},$$

whose columns are the values of T on $\mathcal{B}$.

This reduces the computation of T to the computation of $[\mathbf{x}]_{\mathcal{B}}$.

In simple cases (as above), this can be done by row reduction of the augmented matrix

$$\begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n & \mathbf{x} \end{pmatrix}.$$

A more general approach is to use *matrix inversion*, which we will discuss later.

However, if $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is the standard basis of $\mathbb{R}^n$, then we have

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n \;\Rightarrow\; [\mathbf{x}]_{\mathcal{E}} = \mathbf{x}.$$

Therefore

$$\begin{aligned} T(\mathbf{x}) &= \begin{pmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{pmatrix} [\mathbf{x}]_{\mathcal{E}} \\ &= \begin{pmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{pmatrix} \mathbf{x}. \end{aligned}$$

That is, T is a matrix transformation!

We have therefore proven:

Theorem 5

Every linear transformation is a matrix transformation. Specifically, if $T : \mathbb{R}^n \to \mathbb{R}^m$ *is linear, then* $T(\mathbf{x}) = A\mathbf{x}$ *where*

$$A = \begin{pmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{pmatrix}$$

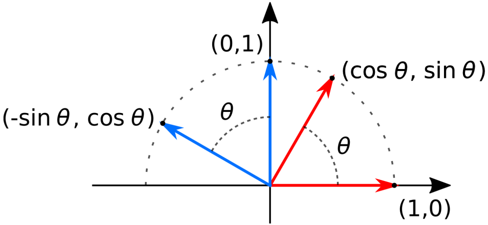
is the $m \times n$ standard matrix *for* T*.*

Let's return to our earlier examples.

Example 4

Find the standard matrix for the linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ given by rotation about the origin by θ radians.

Solution. Rotation of $\mathbf{e}_1$ and $\mathbf{e}_2$ by θ yields:



Therefore the standard matrix is

$$\begin{pmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) \end{pmatrix} = \boxed{\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}},$$

which is called a *rotation matrix*. □

Example 5

Find the standard matrix for the scaling map $T(\mathbf{x}) = a\mathbf{x}$.

Solution. The standard matrix is given by

$$\begin{pmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{pmatrix} = \begin{pmatrix} a\mathbf{e}_1 & a\mathbf{e}_2 & \cdots & a\mathbf{e}_n \end{pmatrix}$$

$$= \boxed{\begin{pmatrix} a & & & \\ & a & & \\ & & \ddots & \\ & & & a \end{pmatrix}},$$

where the blank entries are all equal to 0. □

Remark. Such matrices are called *scalar matrices*.

Example

Let $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be a basis for $\mathbb{R}^n$. How can we find the standard matrix for the coordinate map $[\cdot]_{\mathcal{B}}$?

We need to construct the matrix

$$\begin{pmatrix} [\mathbf{e}_1]_{\mathcal{B}} & [\mathbf{e}_2]_{\mathcal{B}} & \cdots & [\mathbf{e}_n]_{\mathcal{B}} \end{pmatrix}.$$

This amounts to solving the equations

$$\begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{pmatrix} [\mathbf{e}_j]_{\mathcal{B}} = \mathbf{e}_j$$

for $j = 1, 2, \ldots, n$, each of which requires row reduction of the augmented matrix

$$\begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n & \mathbf{e}_j \end{pmatrix}.$$

Because the row operations required to reduce the coefficient matrix are always the same, this can be done *simultaneously* by row reducing

$$\begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n & \mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_n \end{pmatrix},$$

and then reading off the last n columns.

So, if $A = \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{pmatrix}$, $I = \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_n \end{pmatrix}$ (the $n \times n$ *identity matrix*), and B is the standard matrix for $[\cdot]_{\mathcal{B}}$, then

$$\begin{pmatrix} A & I \end{pmatrix} \xrightarrow{\text{RREF}} \begin{pmatrix} I & B \end{pmatrix}.$$

That is, the row operations that transform A into the identity matrix will also transform the identity matrix into the standard matrix for $[\cdot]_{\mathcal{B}}$.

Example 6

If

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \end{pmatrix} \right\},$$

implement the procedure above to find the standard matrix for $[\cdot]_{\mathcal{B}}$.

Solution. We have

$$\begin{pmatrix} 1 & 3 & 1 & 0 \\ 2 & 4 & 0 & 1 \end{pmatrix} \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 0 & -2 & 3/2 \\ 0 & 1 & 1 & -1/2 \end{pmatrix},$$

so that the standard matrix for $[\cdot]_{\mathcal{B}}$ is

$$B = \boxed{\begin{pmatrix} -2 & 3/2 \\ 1 & -1/2 \end{pmatrix}}.$$

□

Remark. Note that this tells us that

$$[\mathbf{x}]_{\mathcal{B}} = B\mathbf{x} = \begin{pmatrix} -2x_1 + \frac{3}{2}x_2 \\ x_1 - \frac{1}{2}x_2 \end{pmatrix},$$

in agreement with our earlier computation.