Matrix-vector products

$A_{21}v_1 + A_{22}v_2 + A_{23}v_3 = w_2$

The simple way: row-by-column.

$$\begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \\ A_{41} & A_{42} & A_{43} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{bmatrix}.$$

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} \to \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix}$$

The smart way: linear combinations of columns of A

$$\begin{bmatrix} A_{11} \\ A_{21} \\ A_{31} \\ A_{41} \end{bmatrix} v_1 + \begin{bmatrix} A_{12} \\ A_{22} \\ A_{32} \\ A_{42} \end{bmatrix} v_2 + \begin{bmatrix} A_{31} \\ A_{32} \\ A_{33} \\ A_{43} \end{bmatrix} v_3 = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{bmatrix}.$$

Im A: set of vectors w that we can obtain.

ker A: possible choices of v that produce zero.

$$\begin{bmatrix} 2 \\ -1 \end{bmatrix} \in \ker \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 0 & 0 \end{bmatrix}$$

Matrix inverses

A

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 9 \\ 15 \end{bmatrix}$$

Linear systems: find coordinates $v_1, \ldots, v_m$ needed to write w as linear combinations of the columns of (square) $A \in \mathbb{R}^{m \times m}$.

A is called invertible if the columns of A generate every vector.

In this case, the solution is given by another matrix: $v = A^{-1} w$

$$AA^{-1} = A^{-1}A = I = \begin{bmatrix} 1 & & & & \\ & 1 & & 0 & \\ 0 & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix}$$

(Convention: omitted elements are zero.)

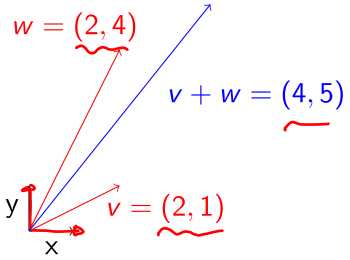
$$A^{-1} \begin{bmatrix} 3 \\ 9 \\ 15 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Bases

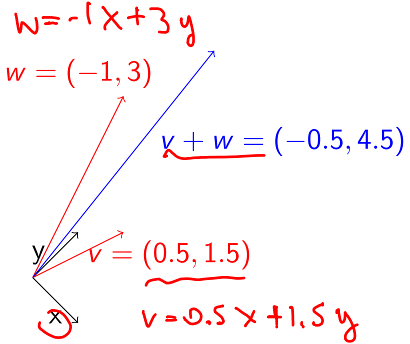
Canonical basis: $w = w_1 e_1 + w_2 e_2 + w_3 e_3 + w_4 e_4$; e.g. for $m = 4$

$$\begin{bmatrix}3\\6\\9\\12\end{bmatrix} = \quad e_1 = \begin{bmatrix}1\\0\\0\\0\end{bmatrix} \cdot 3, \; + \; e_2 = \begin{bmatrix}0\\1\\0\\0\end{bmatrix} \cdot 6, \; + \; e_3 = \begin{bmatrix}0\\0\\1\\0\end{bmatrix} \cdot 9, \; + \; e_4 = \begin{bmatrix}0\\0\\0\\1\end{bmatrix} \cdot 12.$$

Idea behind linear algebra: many things are true regardless of the basis we use.



$\rightarrow$



Matrix-matrix product

$$4\times 3 \quad \begin{array}{r} \\ \\ 3 \\ \\ \end{array}\begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ \boxed{A_{31}} & \boxed{A_{32}} & \boxed{A_{33}} \\ A_{41} & A_{42} & A_{43} \end{bmatrix} \overset{3\times 2}{\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \\ B_{31} & B_{32} \end{bmatrix}} =$$

2

4

$A \in \mathbb{R}^{4\times \boxed{3}}$, $B \in \mathbb{R}^{\boxed{3}\times 2}$. $AB \in \mathbb{R}^{4\times 2}$.
Mnemonic: if the 'inner' dimensions agree we can make the product and 'remove' them.

We can identify vectors with columns ($n \times 1$ matrices).

Cost: multiplying $m \times n$ and $n \times p$ requires $mp(2n-1)$ floating point operations (flops). Forget about fancier algorithms (e.g. Strassen).

Slightly different beast: number-times-matrix, e.g.

$$3A = \begin{bmatrix} 3A_{11} & 3A_{12} & \cdots \\ 3A_{21} & 3A_{22} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}.$$

1×1 $\quad m\times n$

$$\underbrace{A_{31}B_{12} + A_{32}B_{22} + A_{33}B_{32}}_{n}$$

Order of operations

Usual algebra properties hold, e.g.: $A(B+C) = AB + AC$, $A(BC) = (AB)C$, etc.

Warning: Parenthesization matters a lot: if $A, B \in \mathbb{R}^{n \times n}$, $v \in \mathbb{R}^n$, then $(AB)v$ costs $O(n^3)$, but $A(Bv)$ costs $O(n^2)$.

(Matlab example.)
Warning: Matlab does not rearrange parentheses to help you.

$\underbrace{A\overbrace{(Bv)}^{O(n^2)}}_{O(n^2)}$

$\underbrace{(AB)}_{O(n^3)}v$

Matrix algebra

$$(A+B)^2 = (A+B)(A+B) = A^2 + AB + BA + B^2.$$

What doesn't work

$AB \neq BA$: might not even make sense dimension-wise.

Exception: We can move around numbers (scalars): $3AB = A(3B)$.

$AB = AC$ does not imply $B = C$ (example).

However, if there is a matrix M such that $\boxed{MA = I}$, I can multiply by M:

$B = I \cdot B = I \cdot C = C$

$$(MA)B = (MA)C \iff B = C.$$

Warning: multiplying 'on the left' and 'on the right' differ.

(You should remember that for many square matrices there is $M = A^{-1}$ such that $AA^{-1} = A^{-1}A = I$)

(Matlab example: `inv(A)`)

Row and column vectors

$$\underset{3\times 1}{} \quad v = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}, \; v^T = \overset{1\times 3}{\begin{bmatrix} 4 & 5 & 6 \end{bmatrix}}.$$

v is a vector in $\mathbb{R}^3$ (or a matrix in $\mathbb{R}^{3\times 1}$). v^T is a matrix in $\mathbb{R}^{1\times 3}$ (or row vector).

```
>> v = [4;5;6]
v =
     4
     5
     6
>> w = [1 2 3]
w =
     1 2 3
>> w*v
ans =
    32
```

$1\times 3 \;\; 3\times 1 \to 1\times 1$

Row and column vectors

```
>> v*w
ans =
    4  8 12
    5 10 15
    6 12 18
>> v'
ans =
    4  5  6
>> w*v'
Error using *
Inner matrix dimensions must agree.
```

$3 \times 1 \quad 1 \times 3 = 3 \times 3$

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \begin{bmatrix} w_1 & w_2 & w_3 \end{bmatrix} = \begin{bmatrix} v_1 w_1 & v_1 w_2 & v_1 w_3 \\ v_2 w_1 & v_2 w_2 & v_2 w_3 \\ v_3 w_1 & v_3 w_2 & v_3 w_3 \end{bmatrix}$$

Some people are lazy and write *vw* when they mean $v^T w$ — they will burn in hell.

$$(uv)w \neq u(vw)$$

Block operations

The same row-by-column rule holds block-wise.

$$
\begin{array}{cc}
 & \begin{array}{ccc} 3 & 1 & 2 \end{array} \\
\begin{array}{c} 2 \\ \\ 3 \\ \\ 2 \end{array} &
\left[\begin{array}{ccc|c|cc}
* & * & * & * & * & * \\
* & * & * & * & * & * \\ \hline
* & * & * & * & * & * \\
* & * & * & * & * & * \\
* & * & * & * & * & * \\ \hline
* & * & * & * & * & * \\
* & * & * & * & * & *
\end{array}\right]
\end{array}
\cdot
\begin{array}{cc}
 & \begin{array}{ccc} 3 & 2 & 1 \end{array} \\
\begin{array}{c} 3 \\ \\ 1 \\ 2 \end{array} &
\left[\begin{array}{ccc|cc|c}
* & * & * & * & * & * \\
* & * & * & * & * & * \\
* & * & * & * & * & * \\ \hline
* & * & * & * & * & * \\ \hline
* & * & * & * & * & * \\
* & * & * & * & * & *
\end{array}\right]
\end{array}
=
\begin{array}{cc}
 & \begin{array}{ccc} 3 & 2 & 1 \end{array} \\
\begin{array}{c} 2 \\ \\ 3 \\ \\ 2 \end{array} &
\left[\begin{array}{ccc|cc|c}
* & * & * & * & * & * \\
* & * & * & * & * & * \\ \hline
* & * & * & * & * & * \\
* & * & * & * & * & * \\
* & * & * & * & * & * \\ \hline
* & * & * & * & * & * \\
* & * & * & * & * & *
\end{array}\right]
\end{array}
$$

$$
\begin{bmatrix} * & * & * \\ * & * & * \end{bmatrix} \cdot \begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \end{bmatrix} + \begin{bmatrix} * \\ * \end{bmatrix} \cdot \begin{bmatrix} * & * & * \end{bmatrix} + \begin{bmatrix} * & * \\ * & * \end{bmatrix} \cdot \begin{bmatrix} * & * & * \\ * & * & * \end{bmatrix} = \begin{bmatrix} * & * & * \\ * & * & * \end{bmatrix}
$$

In $AB = C$, columns of A and rows of B must be partitioned in the same way, for the product to make sense.

(Matlab example — syntax `A(1:2, 1:3)`.)

Block operations

$A \cdot v_1, \; A \cdot v_2, \; \dots \; A v_n$

$A \left[v_1 \,|\, v_2 \,|\, \cdots \,|\, v_n \right]$

Block operations usually give better performance: one matrix-matrix product performs faster than n matrix-vector products (even if they have the same number of flops). (We are not going to explore it in this course.)

Useful also for analysis: for instance, block triangular matrices:

$$\left[\begin{array}{cc|cc} x & x & x & x \\ x & x & x & x \\ \hline 0 & 0 & x & x \\ 0 & 0 & x & x \end{array}\right]$$

$$\begin{bmatrix} A & B \\ 0 & C \end{bmatrix} \cdot \begin{bmatrix} D & E \\ 0 & F \end{bmatrix} = \begin{bmatrix} AD & AE + BF \\ 0 & CF \end{bmatrix}. \qquad (*)$$

(0 here stands for a block of zeros.)

$$\begin{bmatrix} x & x & x & x \\ 0 & x & x & x \\ 0 & 0 & x & x \\ 0 & 0 & 0 & x \end{bmatrix}$$

Block triangular matrices

Let $\boxed{M} = \begin{bmatrix} \boxed{A_{11}} & A_{12} & \dots & A_{1k} \\ 0 & \boxed{A_{22}} & \dots & A_{2k} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \boxed{A_{kk}} \end{bmatrix}$ be a block triangular matrix, with all A_{ii} square.

- The product of two block (upper/lower) triangular matrices is still triangular — see (*).
- A block triangular matrix is invertible iff all diagonal blocks A_{ii} are invertible
- Its eigenvalues are the union of the eigenvalues of the A_{ii}.

(Matlab example: compute eigenvalues with `eig`).

$$\begin{bmatrix} \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & 0 & \times & \times \\ 0 & 0 & 0 & \times \end{bmatrix} \qquad \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & \times & \times & \times \end{bmatrix}$$

Theorem:

Big matrix invertible $\Longleftrightarrow$ All diagonal blocks are invertible

$$\left[\begin{array}{cc|ccc} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \hline 0 & 0 & \times & \times & \times \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & \times & \times & \times \end{array}\right] \cdot \begin{bmatrix} ? \\ ? \\ ? \\ ? \\ ? \end{bmatrix} = \begin{bmatrix} \times \\ \times \\ \times \\ \times \\ \times \end{bmatrix}$$

Example: 2×2 block triangular linear system

$O(n^3)$

$$\begin{array}{cc} & \begin{array}{cc} n & m \end{array} \\ \begin{array}{c} n \\ m \end{array} & \begin{bmatrix} A & B \\ 0 & C \end{bmatrix} \end{array} \begin{bmatrix} x \\ y \end{bmatrix} \begin{array}{c} n \\ m \end{array} = \begin{bmatrix} e \\ f \end{bmatrix} \begin{array}{c} n \\ m \end{array}.$$

$$\begin{cases} Ax + By = e \\ \boxed{Cy = f} \end{cases}$$

(Again, diagonal blocks are square and all dimensions are compatible.)

$$\begin{bmatrix} Ax + By \\ Cy \end{bmatrix} = \begin{bmatrix} e \\ f \end{bmatrix} \implies y = C^{-1}f, \; x = A^{-1}(e - BC^{-1}f).$$

$$\begin{bmatrix} A & B \\ 0 & C \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} & -A^{-1}BC^{-1} \\ 0 & C^{-1} \end{bmatrix}.$$

(Informal idea: we can start solving from the variables in C.)

$Cy = f$

square matrix

vectors

$$\underbrace{C^{-1}C}_{= I} y = C^{-1} f$$

$$I \cdot y = y$$

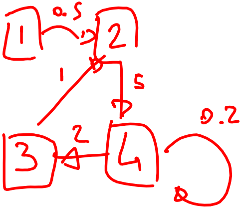
$$y = C^{-1} f$$

$$Ax + By = e$$

$$Ax + BC^{-1} f = e$$

$$Ax = e - BC^{-1} f$$

$$x = A^{-1}\left(e - BC^{-1} f\right)$$



W_{ij}

$$
\begin{array}{c|cccc}
 & 1 & 2 & 3 & 4 \\
\hline
1 & 0 & 0.5 & 0 & 0 \\
2 & 0 & 0 & 0 & 5 \\
3 & 0 & 1 & 0 & 0 \\
4 & 0 & 0 & 2 & 0.2
\end{array}
$$

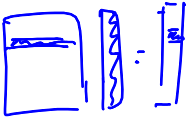
Exercises

1. Write down precisely the dimensions of all matrices in the previous example of a 2×2 block triangular linear system. Be careful — A and C may be square but of different dimensions here, for instance $A \in \mathbb{R}^{m \times m}$, $C \in \mathbb{R}^{n \times n}$.
2. What is the computational cost (up to lower order terms) of computing the product of two square matrices $A, B \in \mathbb{R}^{n \times n}$? Of a matrix-vector product Av, $v \in \mathbb{R}^n$?
3. What is the computational cost of solving a triangular linear system by back-substitution 'starting from the last equation'?
4. Let $A = I + uu^T$, where I is the $n \times n$ identity matrix (what is it?) and u is a vector. How can one compute the product Av (for a vector v) in $O(n)$ flops?

$$A = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix} + \underbrace{\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{bmatrix} \begin{bmatrix} u_1 & u_2 & \cdots & u_m \end{bmatrix}}_{m \times m}$$

$$v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{bmatrix}$$

$$\underset{m\times m}{A} \cdot \underset{m\times 1}{v}$$



m scalar products of length m

$\sim 2m^2$ flops

$$A = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix} + \begin{bmatrix} u \end{bmatrix} \begin{bmatrix} u^T \end{bmatrix}$$

$O(n)$

$O(n^2)$

$$(I + uu^T)v = \underbrace{I \cdot v}_{= v} + u\underbrace{(u^Tv)}_{\text{scalar } 1\times 1}$$

$$= v + (u^Tv)u$$

```
>> k = u'*v;
>> w = v + k*u
```

$\leftarrow 2n-1$

$\leftarrow 2n$

$O(n)$

Exercises

1. Compute the product of two 3×3 block lower triangular matrices, i.e., two of the form
$$\begin{bmatrix} A_{11} & 0 & 0 \\ A_{21} & A_{22} & 0 \\ A_{31} & A_{32} & A_{33} \end{bmatrix}$$
(all A_{ij} here are square matrices, not numbers.) Be careful with the order of the factors.
2. Simplify the expression $A^{-1}(A-B)B^{-1}(A-B)$.
3. What is the inverse of a matrix of the form $\begin{bmatrix} 0 & A \\ B & C \end{bmatrix}$ (all blocks square of the same size)? Is the product of two matrices in this form still in the same form? (Suppose all blocks are square.)
4. Suppose that the adjacency matrix of a graph is block triangular. What does this imply on the graph?abstractname
5. Other exercises (also more challenging) on the Trefethen-Bau and Demmel books.