Similarity transformation and eigenvalue computation

Matrices A and B are similar if there exists an invertible matrix C such that

$$B = C\,A\,C^{-1}$$

$$\downarrow$$

$$A = C^{-1}\,B\,C$$

Suppose $A\,\vec{x} = \lambda\,\vec{x}$

$$\downarrow$$

$$C^{-1}\,B\,C\,\vec{x} = \lambda\,\vec{x}$$

$$\downarrow$$

$$B\,C\,\vec{x} = C\,\lambda\,\vec{x}$$

$$\downarrow$$

$$B\,(C\,\vec{x}) = \lambda\,(C\,\vec{x})$$

$$\downarrow$$

- λ is an eigenvalue of both A and B
- $(C\,\vec{x})$ is an eigenvector of B

$\lambda(A)$ can be computed if we find a C that satisfies

$$T = C\,A\,C^{-1}$$

where T is upper-triangular, lower-triangular, or diagonal,

in which case, $\lambda_i(A) = t_{ii}$

It is best to have C as an orthogonal matrix G, since

$G^T = G^{-1}$, and

$A_{n+1} = G\,A_n\,G^T$ are numerically stable

Therefore, the goal is to find an orthogonal matrix G such that $G\,A\,G^{-1} = T$

Orthogonal matrix G that satisfies

- $A' = G\,A\,G^T$
- $(A')_{2,1} = 0$

Denote a general matrix $A = \begin{bmatrix} d_1 & u \\ \ell & d_2 \end{bmatrix}$

Next, we will try to find G: a matrix of rotation by θ

$$G = \begin{bmatrix} c & s \\ -s & c \end{bmatrix} \quad \text{where } s = \sin(\theta),\ c = \cos(\theta),\ s^2 + c^2 = 1$$

that produces $(A')_{2,1} = 0$ in $A' = G\,A\,G^T$

we do so by expanding

$$G\,A\,G^T = \begin{bmatrix} c & s \\ -s & c \end{bmatrix}\begin{bmatrix} d_1 & u \\ \ell & d_2 \end{bmatrix}\begin{bmatrix} c & -s \\ s & c \end{bmatrix}$$

$$(-s\,d_1 + c\,\ell)\,c + (-s\,u + c\,d_2)\,s = 0$$

$$\downarrow$$

$$\ell\,c^2 - u\,s^2 + (d_2 - d_1)\,c\,s = 0$$

$$\downarrow$$

$$\text{Let } t = \frac{s}{c} = \tan(\theta)$$

$$\downarrow$$

$$\ell - u\,t^2 + (d_2 - d_1)\,t = 0$$

$$\downarrow$$

$$u\,t^2 + (d_1 - d_2)\,t - \ell = 0$$

$$\downarrow$$

$$t = \frac{d_2 - d_1 \pm \sqrt{(d_1 - d_2)^2 + 4\,u\,\ell}}{2\,u}$$

The ± indicates that we may have two $t = \tan(\theta)$ values that satisfy it

suppose $u = \ell$ (A is symmetric) and $u \neq 0$
denote $\Delta = d_2 - d_1$

$$t = \frac{\Delta \pm \sqrt{\Delta^2 + 4\,u^2}}{2\,u}$$ always has two real-valued solutions t_1 and t_2

$$\text{Also, compute } t_1\,t_2 = \frac{\Delta + \sqrt{\Delta^2 + 4\,u^2}}{2\,u} \times \frac{\Delta - \sqrt{\Delta^2 + 4\,u^2}}{2\,u} = -1$$

$$\downarrow$$

θ_1 and θ_2 differ by 90°

since $t_1\,t_2 = -1$, at least one root has $|\,t\,| \leq 1$, corresponding to $|\,\theta\,| \leq 45°$
this is the solution used by the algorithm

For symmetric A,

$$(A')^T = (G\,A\,G^T)^T = (G^T)^T\,A^T\,G^T = G\,A\,G^T = A'$$

$$\downarrow$$

$$(A')_{1,2} = (A')_{2,1} = 0$$

$$\downarrow$$

$$A' = \left[\begin{array}{c|c} d_1' & 0 \\ \hline 0 & d_2' \end{array}\right]$$

Since $A' = G\,A\,G^T$ and G is orthogonal,

$$G^T G = I$$

Multiply $A' = G\,A\,G^T$ on the left by G^T:

$$G^T A' = G^T G\,A\,G^T = A\,G^T$$

Therefore

$$A\,G^T = G^T A'$$

Write $G^T = \left[\,\vec{g}_1 \mid \vec{g}_2\,\right]$

$$\downarrow$$

$$A\left[\,\vec{g}_1 \mid \vec{g}_2\,\right] = \left[\,\vec{g}_1 \mid \vec{g}_2\,\right]\left[\begin{array}{c|c} d_1' & 0 \\ \hline 0 & d_2' \end{array}\right]$$

$$\downarrow$$

$$\left[\, A\vec{g}_1 \mid A\vec{g}_2 \,\right] = \left[\, d_1'\vec{g}_1 \mid d_2'\vec{g}_2 \,\right]$$

$$\downarrow$$

$$A\vec{g}_1 = d_1'\vec{g}_1 \text{ and } A\vec{g}_2 = d_2'\vec{g}_2$$

Therefore d_1' and d_2' are the eigenvalues of symmetric A,

and $\vec{g}_1$ and $\vec{g}_2$ are their corresponding orthonormal eigenvectors.

Two ways to solve the symmetric 2×2 eigenproblem:
Homogeneous systems and Jacobi rotation

Recall how we solved the homogeneous system

$$(A - \lambda I)\,\vec{x} = 0$$

in two stages:

① We first solved for the λ values where the system can have nonzero solutions

② We then substituted each λ and solved the system again
to find those actual solutions: the eigenvectors $\vec{x}$

The Jacobi algorithm gives another route

For a symmetric 2×2 matrix, one similarity transformation is

$$A' = G\,A\,G^{T}$$

For any rotation matrix G: $G^{-1} = G^{T}$

↓

$$A = G^{T} A' G = \begin{bmatrix} c & -s \\ s & c \end{bmatrix} \begin{bmatrix} d_1' & 0 \\ 0 & d_2' \end{bmatrix} \begin{bmatrix} c & s \\ -s & c \end{bmatrix}$$

represents the diagonalization theorem

↓

- $\lambda_1 = d_1'$, $\lambda_2 = d_2'$
- eigenvectors are the orthonormal columns of G^{T}

Step	Homogeneous-system	Jacobi algorithm
Starting equation	$(A - \lambda I)\,\vec{x} = \vec{0}$	$A' = G\,A\,G^T$
First unknown	Eigenvalues λ	Rotation $t = \tan(\theta)$
Condition	$\vec{x} \neq \vec{0}$ $\downarrow$ $\det(A - \lambda I) = 0$	$(A')_{2,1} = 0$
Quadratic equation	$\lambda^2 - (d_1 + d_2)\,\lambda + d_1 d_2 - u^2 = 0$	$u\,t^2 + (d_1 - d_2)\,t - u = 0$
Solutions of the quadratic	$\lambda_{1,2} = \dfrac{d_1 + d_2 \pm \sqrt{(d_1 - d_2)^2 + 4u^2}}{2}$	$t_{1,2} = \dfrac{d_2 - d_1 \pm \sqrt{(d_1 - d_2)^2 + 4u^2}}{2u}$
Eigenvalues	λ_1 and λ_2 are roots of the quadratic eqation	Either value of t produces a diagonal A' The algorithm uses $\lvert t\rvert \leq 1$ $d_1' = d_1 + t\,u$ $d_2' = d_2 - t\,u$ $\{d_1', d_2'\} = \{\lambda_1, \lambda_2\}$ The other root reverses their order
Values computed from the chosen t	Not applicable	$c = \dfrac{1}{\sqrt{1 + t^2}}$, $s = t\,c$
Next calculation	Substitute each λ into $(A - \lambda I)\,\vec{x} = \vec{0}$	Form A' and G^T
Eigenvectors	Solve separately for $\vec{x}_1$ and $\vec{x}_2$	$\vec{g}_1 = (c, s)^T$, $\vec{g}_2 = (-s, c)^T$
Normalization	$\vec{x}_1$ and $\vec{x}_2$ may be normalized after solving	No separate step The columns of G^T already have length 1
Connection to diagonalization theorem	With orthonormal V, $A = V\,D\,V^T$	$V = G^T$ and $D = A'$ Therefore $A = G^T A' G$
Extension to n×n	$\det(A - \lambda I) = 0$ remains true, but its coefficients combine many principal minors and signed products After constructing this degree-n polynomial,	Apply the same 2×2 rotation to any index pair p, q $A_j = G_j\,A_{j-1}\,G_j^T$ After k rotations, let $H = G_k \ldots G_2\,G_1$ $\downarrow$ $A_k = H\,A\,H^T$

a numerical root algorithm
must still converge

This indirect route is not used
by numerical eigensolvers
for large matrices

As A_k converges to diagonal D,
define $Q = H^T$

$$D = Q^T A Q$$
$$A Q = Q D$$

The columns of Q are the
orthonormal eigenvectors of A

How one 2×2 Jacobi rotation
acts inside an n×n matrix

The preceding page showed the 2×2 rotation $R = \begin{bmatrix} c & s \\ -s & c \end{bmatrix}$

that diagonalizes a symmetric 2×2 matrix

We now place that rotation at the nonadjacent
interior coordinates p = 2 and q = 4 of a 5×5 matrix

$$\text{Let } A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{bmatrix}$$

Embed R into rows and columns 2 and 4 of the identity matrix

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & c & 0 & s & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & -s & 0 & c & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \qquad G^T = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & c & 0 & -s & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & s & 0 & c & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

① Right multiplication: $B = A\,G^T$

Columns 1, 3, and 5 of G^T are columns of the identity matrix
Therefore columns 1, 3, and 5 of B are copied unchanged from A

Only columns 2 and 4 of B are recomputed

For every k = 1, ..., 5:

$$b_{k,2} = c\,a_{k,2} + s\,a_{k,4}$$

$$b_{k,4} = -s\,a_{k,2} + c\,a_{k,4}$$

$$B = A\,G^T = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & c & 0 & -s & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & s & 0 & c & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Therefore $B = \begin{bmatrix} a_{1,1} & b_{1,2} & a_{1,3} & b_{1,4} & a_{1,5} \\ a_{1,2} & b_{2,2} & a_{2,3} & b_{2,4} & a_{2,5} \\ a_{1,3} & b_{3,2} & a_{3,3} & b_{3,4} & a_{3,5} \\ a_{1,4} & b_{4,2} & a_{3,4} & b_{4,4} & a_{4,5} \\ a_{1,5} & b_{5,2} & a_{3,5} & b_{5,4} & a_{5,5} \end{bmatrix}$ (Notice $B \neq B^T$)

The colored columns are exactly the entries changed by right multiplication

② Left multiplication: $A' = G\,B$

Rows 1, 3, and 5 of G are rows of the identity matrix

Therefore rows 1, 3, and 5 of B are copied into A′

Only rows 2 and 4 are recomputed

For every $k = 1, \dots, 5$:

$$a'_{2,k} = c\, b_{2,k} + s\, b_{4,k}$$

$$a'_{4,k} = -s\, b_{2,k} + c\, b_{4,k}$$

$$A' = G\,B = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & c & 0 & s & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & -s & 0 & c & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{1,1} & b_{1,2} & a_{1,3} & b_{1,4} & a_{1,5} \\ a_{1,2} & b_{2,2} & a_{2,3} & b_{2,4} & a_{2,5} \\ a_{1,3} & b_{3,2} & a_{3,3} & b_{3,4} & a_{3,5} \\ a_{1,4} & b_{4,2} & a_{3,4} & b_{4,4} & a_{4,5} \\ a_{1,5} & b_{5,2} & a_{3,5} & b_{5,4} & a_{5,5} \end{bmatrix}$$

$$\text{Therefore } A' = \begin{bmatrix} a_{1,1} & a'_{1,2} & a_{1,3} & a'_{1,4} & a_{1,5} \\ a'_{2,1} & d_1' & a'_{2,3} & 0 & a'_{2,5} \\ a_{1,3} & a'_{3,2} & a_{3,3} & a'_{3,4} & a_{3,5} \\ a'_{4,1} & 0 & a'_{4,3} & d_2' & a'_{4,5} \\ a_{1,5} & a'_{5,2} & a_{3,5} & a'_{5,4} & a_{5,5} \end{bmatrix} \qquad \left(A' = (A')^T\right)$$

The two multiplications have now accounted for the complete colored cross:

- right multiplication changed columns 2 and 4
- left multiplication changed rows 2 and 4
- if $i, j \neq 2, 4$, then $a'_{i,j} = a_{i,j}$

The four entries at the intersection of these rows and columns undergo exactly the 2×2 multiplication derived previously

$$\text{Active block of } A' = \begin{bmatrix} d_1' & 0 \\ 0 & d_2' \end{bmatrix} = \begin{bmatrix} c & s \\ -s & c \end{bmatrix} \begin{bmatrix} d_1 & u \\ u & d_2 \end{bmatrix} \begin{bmatrix} c & -s \\ s & c \end{bmatrix}$$

Thus the 2×2 equation is solved once, and c and s are inserted into G

- Right-multiplication updates the two active columns
- Left-multiplication updates the two active rows
- The same 2×2 calculation diagonalizes their intersection
- All the entries shown in black remain unchanged, facilitating computations

This construction generalizes to any symmetric n×n matrix

Why not systematically zero all subdiagonal entries?

A later similarity transformation can bring a zero back to life

$$\text{Suppose } A' = \begin{bmatrix} a'_{11} & a'_{12} & a'_{13} \\ 0 & a'_{22} & a'_{23} \\ a'_{31} & a'_{32} & a'_{33} \end{bmatrix} \qquad H = \begin{bmatrix} 1 & 0 & 0 \\ 0 & c & s \\ 0 & -s & c \end{bmatrix}$$

Indices: $p = 1$, $q = 2$, $r = 3$

- Green: $(A')_{qp} = 0$ is the subdiagonal zero we earned
- Red: $(A')_{rp}$ is a nonzero entry below it in the same column

$$\text{col } p(B = A' H^T) = \text{col } p(A') = \begin{bmatrix} a'_{11} \\ 0 \\ a'_{31} \end{bmatrix}$$

Now $A'' = H B$

$\downarrow$

$$(A'')_{qp} = \text{row } q(H) \times \text{col } p(B) = \begin{bmatrix} 0 & c & s \end{bmatrix} \begin{bmatrix} a'_{11} \\ 0 \\ a'_{31} \end{bmatrix} = s\, a'_{31}$$

$\downarrow$

For a nontrivial rotation, $s\, a'_{31}$ is generally nonzero

So the corresponding p, q, r block of $A'' = H A' H^T$ becomes
$$\begin{bmatrix} a''_{11} & a''_{12} & a''_{13} \\ s\, a'_{31} & a''_{22} & a''_{23} \\ c\, a'_{31} & a''_{32} & a''_{33} \end{bmatrix}$$

- A previous similarity transformation created a subdiagonal zero
- A later similarity transformation filled it back in

Why does the Jacobi algorithm converge?

Even though a previously created zero may be filled back in,
repeated similarity transformations for symmetric A
eventually converge to a diagonal matrix

Next, we define two quantities:

① Off-diagonal energy, $\Omega^2(A)$: the sum of squared entries a_{ij} with $i \neq j$:

$$\Omega^2(A) = 2 \sum_{i=1}^{n-1} \left(\sum_{j=i+1}^{n} \left(a_{ij}\right)^2 \right)$$

② Frobenius norm, $\|A\|_F$, whose square is the sum of squared entries of A:

$$\left(\|A\|_F\right)^2 = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} \left(a_{ij}\right)^2 \right)$$

Therefore

$$\left(\|A\|_F\right)^2 = \sum_{i=1}^{n} \left(a_{ii}{}^2\right) + \Omega^2(A)$$

For the matrix $A = A^T = \begin{bmatrix} d_1 & u & a_{13} \\ u & d_2 & a_{23} \\ a_{13} & a_{23} & d_3 \end{bmatrix}$

$$\Omega^2(A) = 2\,(u^2 + a_{13}^{\ 2} + a_{23}^{\ 2})$$

After applying a Jacobi rotation to the pivot $u = a_{12}$, we obtain:

$$A' = G\,A\,G^T = \begin{bmatrix} d_1' & 0 & a'_{13} \\ 0 & d_2' & a'_{23} \\ a'_{13} & a'_{23} & d_3 \end{bmatrix}$$

G is orthogonal

↓

- left-multiplication by G preserves length of every column of A
- right-multiplication by G^T preserves length of every row of A

↓

The similarity transformation $G\,A\,G^T$ preserves the Frobenius norm:

$$\|A'\|_F^2 = \|A\|_F^2$$

Thus the total squared magnitude of A does not change.

The active 2×2 block rotates from $\begin{bmatrix} d_1 & u \\ u & d_2 \end{bmatrix}$ to $\begin{bmatrix} d_1' & 0 \\ 0 & d_2' \end{bmatrix}$

↓

Its total squared magnitude is also preserved:

$$d_1'^2 + d_2'^2 = d_1^2 + d_2^2 + 2u^2$$

↓

Thus the diagonal $\sum_{i=1}^{n} \left(a_{ii}^{2}\right)$ gains $2u^2$ of squared magnitude.

Since the total squared magnitude of A is unchanged,
that same $2u^2$ must be subtracted from the off-diagonal energy

↓

$$\Omega^2(A') = \Omega^2(A) - 2u^2$$

More generally, for any pivot $a_{p,q}$:

$$\Omega^2(A') = \Omega^2(A) - 2\left(a_{p,q}\right)^2$$

↓

Every nonzero Jacobi rotation strictly decreases the off-diagonal energy

Since the largest off-diagonal pivot always contains a fixed fraction of the remaining off-diagonal energy, repeated Jacobi rotations force

$\Omega^2(A) \rightarrow 0$, and A converges to a diagonal matrix.

Single Jacobi rotation in 2D

Consider an example of $A = A^{T} = \begin{bmatrix} 1.8 & 0.1 \\ 0.1 & 0.8 \end{bmatrix}$

For a 2×2 symmetric matrix, the algorithm uses one similarity transformation:

G = rotation by ≈-6° = $\begin{bmatrix} \approx 0.995 & \approx 0.099 \\ \approx -0.099 & \approx 0.995 \end{bmatrix}$ $D = G A G^{T} = \begin{bmatrix} \approx 1.81 & 0 \\ 0 & \approx 0.79 \end{bmatrix}$

Ⓐ

Jacobi update: $D = G A G^{T}$

① left: transformation by $A = A^{T}$

↓

for $\lambda_1 \neq \lambda_2$, the corresponding eigenvectors $\vec{g}_1$ and $\vec{g}_2$ are orthogonal and $G^{T} = [\, \vec{g}_1 \mid \vec{g}_2 \,]$

② middle: right-multiplication by G^{T}

$$A G^{T} = A [\, \vec{g}_1 \mid \vec{g}_2 \,]$$
$$= [\, A \vec{g}_1 \mid A \vec{g}_2 \,]$$
$$= [\, \lambda_1 \vec{g}_1 \mid \lambda_2 \vec{g}_2 \,]$$

Right-multiplication by a rotation matrix is NOT a rotation of the displayed transformation

it forms new columns from linear combinations of the columns of A

and preserves lengths and angles between rows, not between columns

Here those new columns are $\lambda_1 \vec{g}_1$ and $\lambda_2 \vec{g}_2$

so they point along the two orthogonal eigenvector directions

③ right: left-multiplication by G

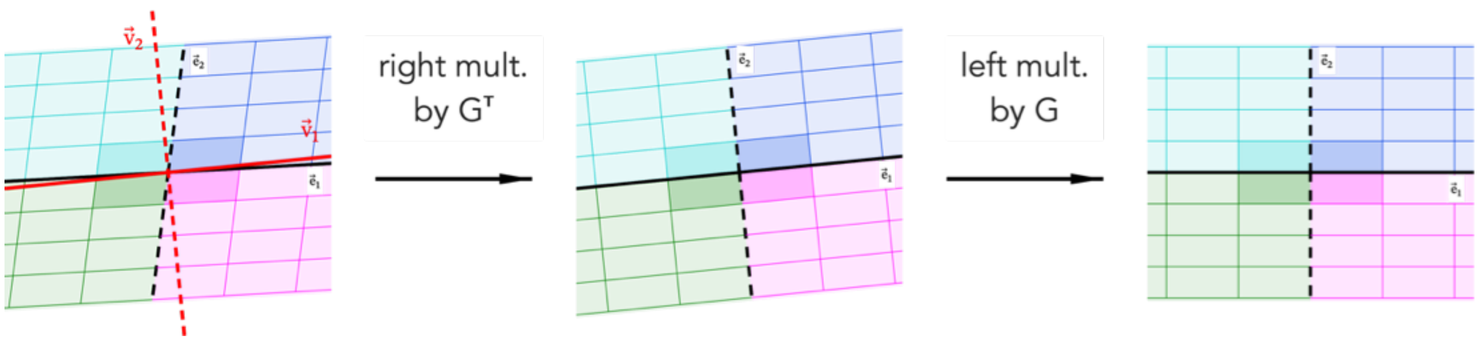
left-multiplication applies the same rotation G to every column

so this step truly rotates the displayed vectors

$$G\,\vec{g}_1 = \vec{e}_1 \text{ and } G\,\vec{g}_2 = \vec{e}_2$$

$$\downarrow$$

$$G\,A\,G^{T} = [\ \lambda_1\,\vec{e}_1 \mid \lambda_2\,\vec{e}_2\ ] = D$$



Ⓑ

The same Jacobi update with the transformation of the unit circle added

• Notice that the ellipse remains unchanged after right-multiplication

Let $C = \{\ \vec{x} : \vec{x}^{T}\,\vec{x} = 1\ \}$ be the set of vectors on the unit circle

since G^{T} is orthogonal,

$$(G^{T}\,\vec{x})^{T}(G^{T}\,\vec{x}) = \vec{x}^{T}\,G\,G^{T}\,\vec{x} = \vec{x}^{T}\,\vec{x}$$

$$\downarrow$$

$$G^{T}\,C = C$$

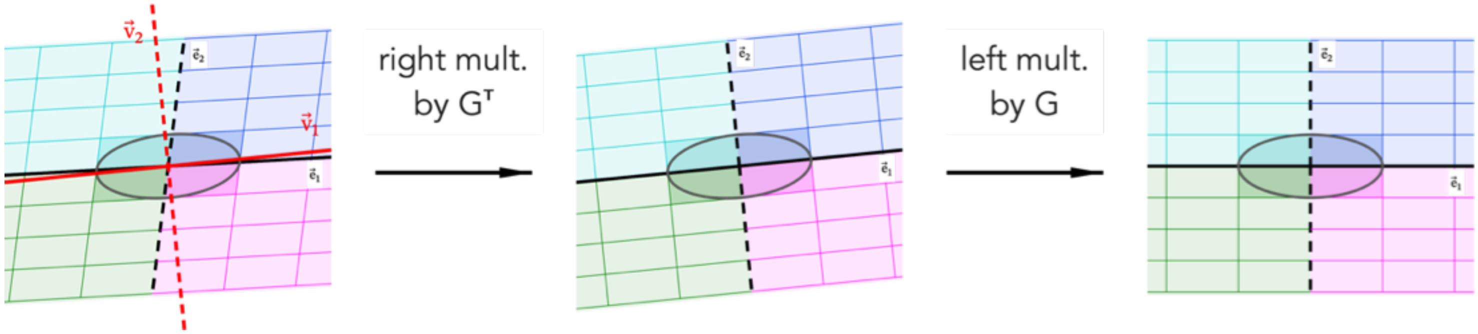
$$(A\,G^{T})(C) = A(G^{T}\,C) = A(C)$$

Thus A and $A\,G^{T}$ map the unit circle to exactly the same ellipse

• Left-multiplication is different:

$$(G\,A\,G^{T})(C) = G\big(A(C)\big)$$

so G now rotates the ellipse itself



©

Same transformation shown as factorization of A or diagonalization

$A = G^T D G$