Second Half of the Linear Algebra Course (**wi1 277LR**)

Determinants

Cofactor expansion across the ith row: $\det A = a_{i1}C_{i1} + \cdots + a_{in}C_{in}$ with cofactor $C_{ij} = (-1)^{i+j} \det A_{ij}$. Cofactor expansion across an arbitrary row or column of A gives the same answer. Then also: $\det A^T = \det A$.

Row operations $A \rightsquigarrow B$:
row replacement $\Rightarrow \det A = \det B$, row interchange $\Rightarrow \det B = -\det A$, row multiplication by $k \Rightarrow \det B = k \det A$, and therefore $\det(kA) = k^n \det A$.
$\Rightarrow$ First simplify the determinant by row or column operations before expanding!

Properties:
$$A^{-1} \text{ exists} \quad \Leftrightarrow \quad \det A \neq 0$$
$$\det(AB) \quad = \quad (\det A)(\det B), \quad \text{therefore: } \det(A^{-1}) = 1/\det A$$

Area of parallelogram, volume of parallelepiped:
2×2 matrix $A = [\mathbf{a}_1\, \mathbf{a}_2] \;\Rightarrow\;$ area$\{\mathbf{a}_1, \mathbf{a}_2\}$ = absolute value of $\det A$.
3×3 matrix $A = [\mathbf{a}_1\, \mathbf{a}_2\, \mathbf{a}_3] \;\Rightarrow\;$ volume$\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ = absolute value of $\det A$.

If S is a region of $\mathbb{R}^k$ and $T : \mathbb{R}^k \to \mathbb{R}^k$ is a linear transformation with $T(\mathbf{x}) = A\mathbf{x}$ then: volume of $T(S) = |\det A| \cdot$ volume of S.

Theoretical properties

- $T(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n) = \det[\mathbf{x}_1\, \mathbf{x}_2 \cdots \mathbf{x}_n]$ is a multi linear function, that is: linear in each of its arguments[1].
- **Cramer's rule** Let A be invertible and $A\mathbf{x} = \mathbf{b}$. The matrix $A_i(\mathbf{b})$ is obtained by replacing column i by the vector $\mathbf{b}$. Then $x_i = \frac{\det A_i(\mathbf{b})}{\det A}$.
- A formula for A^{-1} by means of a matrix of cofactors, called the adjugate of A.

Vector space

A vector space is a nonempty set V of objects, called vectors, closed under two operations: addition and multiplication by scalars, and satisfying certain rules.
The set contains a zero vector $\mathbf{0}$ with $\mathbf{u} + \mathbf{0} = \mathbf{u}$, and $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$.

- $\mathbb{R}^n$. The zero vector is $\mathbf{0} = [0\,0 \cdots 0]^T$.
- The set of all real-valued functions $f : D \to \mathbb{R}$, with the well known addition and scalar multiplication for functions. Zero function $f(t) = 0$ for all $t \in D$ acts as the zero vector.
- $\mathbb{P}_n$ (set of polynomials of degree at most n).
The zero polynomial [$\mathbf{p}(t) = 0$ for all t] acts as the zero vector.

---

[1] §3.2 p.197 Proof by expanding across the appropriate column.

One can apply *earlier notions*: linear combination of vectors, linearly independent set of vectors, span of a set of vectors, subspace, basis, dimension, coordinate vector.

Example: a (standard) basis of $\mathbb{P}_n$ is the collections of polynomials $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_{n+1}\}$ with $\mathbf{e}_1(t) = 1,\ \mathbf{e}_2(t) = t,\ \mathbf{e}_3(t) = t^2,\ \cdots, \mathbf{e}_{n+1}(t) = t^n$, and so $\dim \mathbb{P}_n = n + 1$.

If $\mathbf{p}$ is a polynomial of $\mathbb{P}_2$ with $\mathbf{p}(t) = -2+t+3t^2$ then $[\mathbf{p}]_{\mathcal{E}} = \begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix}$.

Coordinate systems

Let $\mathcal{B} = \{\mathbf{b}_1, \cdots, \mathbf{b}_n\}$ be a basis for vector space V. The coordinate mapping $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$ is a one-to-one linear transformation from V onto $\mathbb{R}^n$. Linearity of the mapping:

$$[c\,\mathbf{u} + d\,\mathbf{v}]_{\mathcal{B}} = c\,[\mathbf{u}]_{\mathcal{B}} + d\,[\mathbf{v}]_{\mathcal{B}}$$

Special case: $V = \mathbb{R}^n$. Since $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$ and $\mathbf{x} = c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n$, therefore $\mathbf{x} = P_{\mathcal{B}}\,[\mathbf{x}]_{\mathcal{B}}$ with the change-of-coordinates matrix $P_{\mathcal{B}} = [\mathbf{b}_1\,\mathbf{b}_2 \cdots \mathbf{b}_n]$.

The matrix representation of linear transformations

The information about vectors can be stored in a column [coordinate vector]. *The information about a* **linear** *map (transformation) can be stored in a matrix*[2].

Let V be a an n-dimensional vector space, W an m-dimensional vectorspace, and $T : V \to W$ a linear transformation from V tot W, that is:

$T(c\,\mathbf{u} + d\,\mathbf{v}) = c\,T(\mathbf{u}) + d\,T(\mathbf{v})$ for all $\mathbf{u}$, $\mathbf{v}$ from V and all scalars c,d

Choose basis $\mathcal{B} = \{\mathbf{b}_1, \cdots, \mathbf{b}_n\}$ for V and basis $\mathcal{C}$ for W. Then

$$[T(\mathbf{x})]_{\mathcal{C}} \;=\; M\,[\mathbf{x}]_{\mathcal{B}} \quad \text{where } M = [\,[T(\mathbf{b}_1)]_{\mathcal{C}}\,[T(\mathbf{b}_2)]_{\mathcal{C}}\,\ldots[T(\mathbf{b}_n)]_{\mathcal{C}}\,]$$

M is called a matrix representation of T relative to bases $\mathcal{B}$ and $\mathcal{C}$.

Special case: Linear transformation T from V to V with basis $\mathcal{B}$ for V. The $\mathcal{B}$-matrix for T is $[T]_{\mathcal{B}} = [\,[T(\mathbf{b}_1)]_{\mathcal{B}}\,[T(\mathbf{b}_2)]_{\mathcal{B}}\,\ldots[T(\mathbf{b}_n)]_{\mathcal{B}}\,]$.

Example Consider the linear map $D : \mathbb{P}_2 \to \mathbb{P}_2$ with $D(\mathbf{p}) = \mathbf{p}'$ (the derivative). Standard basis $\mathcal{E} = \{1, t, t^2\}$ for $\mathbb{P}_2$. Then the $\mathcal{E}$-matrix of D is

$$[D]_{\mathcal{E}} \;=\; \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

[2] See §5.4.

Similarity

Square matrices A and B are similar[3] if there is an invertible P with $B = P^{-1}AP$.

Similarity of Matrix Representations[4]
Given a basis $\mathcal{B} = \{\mathbf{b}_1, \cdots, \mathbf{b}_n\}$ for $\mathbb{R}^n$ and a matrix transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ with $T(\mathbf{x}) = A\mathbf{x}$. Then $[T]_{\mathcal{B}} = PAP^{-1} = B$ with $P = [\mathbf{b}_1 \cdots \mathbf{b}_n]$.

Application
Consider the discrete dynamical system $\mathbf{x}_0 \xrightarrow{A} \mathbf{x}_1 \xrightarrow{A} \mathbf{x}_2 \xrightarrow{A} \cdots$ [the difference equation $\mathbf{x}_{k+1} = A\mathbf{x}_k$]. Let be given the change-of-coordinates matrix $P = P_{\mathcal{B}} = [\mathbf{b}_1 \cdots \mathbf{b}_n]$.
If $\mathbf{x}_k = P\mathbf{y}_k$ and therefore $[\mathbf{x}_k]_{\mathcal{B}} = \mathbf{y}_k$, then $\mathbf{y}_0 \xrightarrow{B} \mathbf{y}_1 \xrightarrow{B} \mathbf{y}_2 \xrightarrow{B} \cdots$ with $B = PAP^{-1}$.

Eigenvalues and eigenvectors Given: A is a $n \times n$-matrix.
$\mathbf{v}$ is an eigenvector of the square matrix A if $\mathbf{v} \neq \mathbf{0}$ and $A\mathbf{v} = \lambda\mathbf{v}$ for certain λ. The scalar λ is then called an eigenvalue of A and $E_\lambda \equiv \{\mathbf{x} \mid A\mathbf{x} = \lambda\mathbf{x}\} = \mathsf{Nul}(A - \lambda I)$ is called the eigenspace of A corresponding to λ. Remark: $\dim E_\lambda \geq 1$.

Geometric example: Let A be the matrix of the orthogonal projection $\mathbf{x} \mapsto A\mathbf{x}$ in $\mathbb{R}^3$ on a plane W through $\mathcal{O}$. Then A has two eigenspaces: $E_1 = W$ and $E_0 = W^\perp$ (= line perpendicular to W through $\mathcal{O}$).
Discrete dynamical system: $\mathbf{x}_0 \xrightarrow{A} \mathbf{x}_1 \xrightarrow{A} \mathbf{x}_2 \xrightarrow{A} \cdots$. If $\mathbf{x}_0$ is an eigenvector of A with eigenvalue λ then $\mathbf{x}_k = \lambda^k\mathbf{x}_0$.

- A (real) matrix A may not have (real) eigenvalues.
- If $\mathbf{v}_1, \cdots, \mathbf{v}_p$ are eigenvectors corresponding to *distinct* eigenvalues, then
  the set $\{\mathbf{v}_1, \cdots, \mathbf{v}_p\}$ is linearly independent.

Calculating all the eigenvalues and eigenspaces
First solve the *characteristic equation* $\det(A - \lambda I) = 0$. Then for each solution λ solve $(A - \lambda I)\mathbf{x} = \mathbf{0}$, that is: determine a basis for each E_λ.

- $\det(A - \lambda I)$ is a polynomial in λ of degree n (characteristic polynomial).
- The (algebraic) multiplicity α_λ of an eigenvalue λ is its multiplicity as a root of the characteristic equation. $1 \leq \alpha_\lambda \leq n$.

[3] See §5.2. Remark: 'similar' (='gelijkvormig') , do not confuse with 'row equivalent' (='rij-equivalent')
[4] See §5.4.

- Theorem: $\text{dim}(E_\lambda) \leq \alpha_\lambda$. That is:
  The dimension of the eigenspace E_λ, also called the geometric multiplicity of the eigenvalue λ, is less than or equal to the algebraic multiplicity of λ. ([5])
- 0 is an eigenvalue $\Leftrightarrow$ A is not invertible
- Similar matrices have the same characteristic polynomial, therefore the same eigenvalues λ_i with the same multiplicities α_{λ_i}.

Diagonalization

A is diagonalizable if A is similar to a diagonal matrix, that is: there is an invertible matrix P and a diagonal matrix D such that $A = PDP^{-1}$.

Theorem The $n \times n$ matrix A is diagonalizable $\Leftrightarrow$ there is a basis of $\mathbb{R}^n$ consisting of eigenvectors of A.

In the diagonal of D: eigenvalues of A. The columns of P form a (corresponding) basis of eigenvectors. $D = \text{diag}(\lambda_1, \cdots, \lambda_n)$ and $P = [\mathbf{p}_1\, \mathbf{p}_2 \cdots \mathbf{p}_n]$ with $A\mathbf{p}_i = \lambda_i\, \mathbf{p}_i$.

Theorem Let A be an $n \times n$ matrix with distinct eigenvalues $\lambda_1, \cdots, \lambda_p$.

- A is diagonalizable $\Leftrightarrow$ $\displaystyle\sum_{i=1}^{p} \text{dim} E_{\lambda_i} = n$
- A is diagonalizable $\Leftrightarrow$ $\begin{cases} \text{for all eigenvalues } \lambda: \ \text{dim} E_\lambda = \alpha_\lambda \\ \displaystyle\sum_{i=1}^{p} \alpha_{\lambda_i} = n \quad \text{("enough eigenvalues")} \end{cases}$

Calculation of D and P
Determine for each eigenvalue λ_k a basis $\mathcal{B}_k$ for the eigenspace E_{λ_k}. If the total collection of vectors in $\mathcal{B}_1, \cdots, \mathcal{B}_p$ has n vectors, then A is diagonalizable and this collection forms an eigenvector basis for $\mathbb{R}^n$. Put this basis in P and make the diagonal matrix D with the corresponding eigenvalues in the diagonal.

Diagonal Matrix Representation
If A is diagonalizable: $A = PDP^{-1}$, then the transformation $\mathbf{x} \mapsto T(\mathbf{x}) = A\mathbf{x}$ has a very simple matrix representation, namely by a diagonal matrix: $[T]_\mathcal{B} = D$ where $\mathcal{B}$ is the basis for $\mathbb{R}^n$ formed from the columns of P.

Complex eigenvalues

The characteristic equation of the square matrix A has always exactly n roots, *provided that possibly complex roots are included.*

[5] The matrix A is called defective if $\text{dim} E_\lambda < \alpha_\lambda$ for some eigenvalue λ.

This complex root λ is an eigenvalue of A when we let A act on the space $\mathbb{C}^n$ of n-tuples of complex numbers: $A\mathbf{x} = \lambda\mathbf{x}$ with $\mathbf{x}$ in $\mathbb{C}^n$. ([6])

If $\mathbf{x}$ from $\mathbb{C}^n$ then we can form: $\mathrm{Re}\,\mathbf{x}$, $\mathrm{Im}\,\mathbf{x}$, the complex conjugate $\overline{\mathbf{x}}$.
When A is a real matrix then complex eigenvalues occur in conjugate pairs, that is:
if $A\mathbf{x} = \lambda\,\mathbf{x}$ and A is a real matrix, then $A\,\overline{\mathbf{x}} = \overline{\lambda}\,\overline{\mathbf{x}}$.

Example Let $C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ with real a, $b \neq 0$.
Then C is a rotation followed by a scaling: $C = r \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}$, with ϕ and r being the argument and modulus of the complex eigenvalue $\lambda = a + b\,i$, that is: $\lambda = a + bi = r\,e^{i\phi}$.

Any 2×2 matrix with complex eigenvalue is similar to a rotation followed by scaling. Theorem Let A be a real 2×2 matrix with a complex eigenvalue $\mu = a - b\,i$ $(b \neq 0)$ and associated eigenvector $\mathbf{v}$ in $\mathbb{C}^2$. Then $A = PCP^{-1}$ where $P = [\mathrm{Re}\,\mathbf{v} \;\; \mathrm{Im}\,\mathbf{v}]$ and $C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$.
For the rotation angle and scaling factor: see the example.

Symmetric matrices

A matrix A is orthogonally diagonalizable if there is an orthogonal matrix P (that is: $P^{-1} = P^T$) and a diagonal matrix D such that $A = PDP^{-1}$.

A symmetric matrix is a matrix A such that $A^T = A$. For a symmetric matrix any two eigenvectors from *different* eigenspaces are orthogonal. The eigenspaces are therefore mutually orthogonal.

- If A is orthogonally diagonalizable, then A is symmetric (simple proof).
- *Proposition*[7]: If A is symmetric then A is diagonalizable.

By choosing/constructing an orthonormal basis for each eigenspace, one gets an orthonormal basis of eigenvectors for the whole space and from this basis an orthogonal matrix P which does diagonalize matrix A. Therefore:

Theorem A is symmetric $\Leftrightarrow$ A is orthogonally diagonalizable.

[6] The dimension of the complex eigenspace may be less than the algebraic multiplicity of the eigenvalue (the matrix A is then called defective). Otherwise the matrix is complex diagonalizable.

[7] Its proof is not simple.

Spectral decomposition If $A = PDP^{-1}$ with orthogonal matrix $P = [\mathbf{u}_1 \cdots \mathbf{u}_n]$ and $D = \mathsf{diag}(\lambda_1, \ldots, \lambda_n)$, then $A = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \cdots \lambda_n \mathbf{u}_n \mathbf{u}_n^T$.

The Principle Axes Theorem for quadratic forms

A quadratic form $Q : \mathbb{R}^n \to \mathbb{R}$ is defined by $Q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$ where A is a symmetric matrix. Let P orthogonally diagonalize A with $A = PDP^T$. Then a change of variable $\mathbf{x} = P\mathbf{y}$ gives $\mathbf{x}^T A \mathbf{x} = \mathbf{y}^T D \mathbf{y}$, an expression for the quadratic form without cross-product terms. The columns of P are the principal axes of the quadratic form Q.

See: a geometric view of principal axes. Graph of the level set $Q(\mathbf{x}) = \mathsf{constant}$ in standard position, principal axes determined by the orthonormal eigenvectors.

Application: three-dimensional dynamics of rigid bodies. The rotation energy of a rigid body is a quadratic form based on the symmetric inertia matrix.

Classification of a quadratic form $Q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$

A quadratic form Q is positive definite if $Q(\mathbf{x}) > 0$ for all $\mathbf{x} \neq \mathbf{0}$.
A quadratic form Q is positive semidefinite if $Q(\mathbf{x}) \geq 0$ for all $\mathbf{x}$.
Q is indefinite if Q assumes both positive and negative values.

Theorem The quadratic form Q is positive definite $\Leftrightarrow$ eigenvalues of A are all positive. Q is indefinite $\Leftrightarrow$ A has both positive and negative eigenvalues (*proof* by orthogonally diagonalizing).

APPLICATION **Discrete Dynamical Systems $\mathbf{x}_{k+1} = A\,\mathbf{x}_k$**

Consider $\mathbf{x}_0 \xrightarrow{A} \mathbf{x}_1 \xrightarrow{A} \mathbf{x}_2 \xrightarrow{A} \cdots$ with diagonalizable A. Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be a basis of eigenvectors with corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$.

The general solution of the system $\mathbf{x}_{k+1} = A\,\mathbf{x}_k$ can be written as

$$\mathbf{x}_k = c_1\,(\lambda_1)^k\,\mathbf{v}_1 + \cdots + c_n\,(\lambda_n)^k\,\mathbf{v}_n \quad \text{for arbitrary scalars } c_1, \ldots, c_n$$

The eigenvector decomposition of $\mathbf{x}_0$ determines what happens in $\mathbf{x}_0 \xrightarrow{A} \mathbf{x}_1 \xrightarrow{A} \cdots$ since the coefficients c_i are determined by $\mathbf{x}_0 = c_1\,\mathbf{v}_1 + \cdots + c_n\,\mathbf{v}_n$.

Long term behavior

- The system $\mathbf{x}_{k+1} = A\,\mathbf{x}_k$ has infinitely many solutions, one of them is the (trivial) constant zero solution: $\mathbf{x}_k = \mathbf{0}$ for all k.
- If $|\lambda_1| \geq 1 > |\lambda_j|$ for $j \neq 1$, then for sufficiently large k: $\mathbf{x}_{k+1} \approx c_1\,(\lambda_1)^k\,\mathbf{v}_1$.
- $\mathbf{0}$ is called an **attractor** if for all eigenvalues $|\lambda_i| < 1$. All trajectories of $\mathbf{x}_{k+1} = A\,\mathbf{x}_k$ tend toward $\mathbf{0}$.
- $\mathbf{0}$ is called a **repellor** if for all eigenvalues $|\lambda_i| > 1$. All trajectories of $\mathbf{x}_{k+1} = A\,\mathbf{x}_k$ (except the constant zero solution) tend away from the origin.
- $\mathbf{0}$ is called a **saddle point** if for some eigenvalues $|\lambda_i| > 1$ and for the other eigenvalues $|\lambda_j| < 1$. The origin attracts solutions from some directions and repels them in other directions.

Change of variables

Let $A = PDP^{-1}$. Consider $\mathbf{x}_0 \xrightarrow{A} \mathbf{x}_1 \xrightarrow{A} \mathbf{x}_2 \xrightarrow{A} \cdots$, take P as the change-of-coordinates matrix and define $\mathbf{y}_k = P^{-1}\mathbf{x}_k$. Then $\mathbf{y}_0 \xrightarrow{D} \mathbf{y}_1 \xrightarrow{D} \mathbf{y}_2 \xrightarrow{D} \cdots$ with D a diagonal matrix and therefore one has in these new variables a very simple discrete system $\mathbf{y}_{k+1} = D\mathbf{y}_k$. The system is in these new variables decoupled.

Graphical picture (in case $n = 2$) With $P = [\mathbf{v}_1 \ \mathbf{v}_2]$, draw axes from the origin through $\mathbf{v}_1$ and $\mathbf{v}_2$ and make a graph with trajectories as viewed in terms of these eigenvector axes.

Discrete Dynamical System with complex eigenvalues (in case $n = 2$)

If real A has two complex eigenvalues (λ and $\overline{\lambda}$) whose absolute value is greater than 1, then $\mathbf{0}$ is a repellor: the trajectory spirals outward around the origin. If the absolute values are less than 1, the origin is an attractor (inwards spiralling trajectories).

APPLICATION **Continuous Dynamical Systems $\mathbf{x}'(t) = A\mathbf{x}(t)$**

Consider a continuous dynamical system described by a linear system of differential equations (of first order) $\mathbf{x}'(t) = A\,\mathbf{x}(t)$ with diagonalizable matrix A.

Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be a basis of eigenvectors with corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$, that is: $A = PDP^{-1}$ with $P = [\mathbf{v}_1 \cdots \mathbf{v}_n]$, $D = \mathsf{diag}(\lambda_1, \ldots, \lambda_n)$.

- The constant solution $\mathbf{x}'(t) = \mathbf{0}$ for all t, is the *trivial* solution of a system $\mathbf{x}'(t) = A\mathbf{x}(t)$.
- Fundamental set of eigenfunctions: $\mathbf{v}_1\,e^{\lambda_1 t}, \cdots, \mathbf{v}_n e^{\lambda_n t}$. These are (basic) solutions of the system $\mathbf{x}'(t) = A\,\mathbf{x}(t)$.
- The general solution of the system is

$$\mathbf{x}(t) = c_1\mathbf{v}_1\,e^{\lambda_1 t} + \cdots c_n\mathbf{v}_n\,e^{\lambda_n t}$$

(proof by *decoupling* the system through a change of variables $\mathbf{x} = P\mathbf{y}$)

The eigenvector decomposition of $\mathbf{x}(0) = \mathbf{x}_0$ (initial value) determines the value of the coefficients c_i.

Long term behavior

- If all eigenvalues $\lambda > 0$, then the non trivial solutions tend away from $\mathbf{0}$ as $t \to \infty$, since in this case $e^{\lambda t} \to \infty$. The eigenvectors belonging to the greatest eigenvalue give the direction of greatest repulsion. $\mathbf{0}$ is a repellor.
- If If all eigenvalues $\lambda < 0$, then the non trivial solutions tend to $\mathbf{0}$ as $t \to \infty$, since in this case $e^{\lambda t} \to 0$. $\mathbf{0}$ is an attractor. The eigenvectors belonging to the most negative eigenvalue give the direction of greatest attraction.
- If some eigenvalues are positive and the other eigenvalues are negative then $\mathbf{0}$ is a saddle point.

Complex eigenvalues (in case $n = 2$)

The real 2×2 matrix A has a pair of complex eigenvalues $\lambda = a + bi$ and $\overline{\lambda} = a - bi$ ($b \neq 0$), with associated complex eigenvectors $\mathbf{v}$ and $\overline{\mathbf{v}}$.

A general *complex* solution is

$$\mathbf{x}(t) = d_1\,\mathbf{v}e^{\lambda t} + d_2\,\overline{\mathbf{v}}e^{\overline{\lambda} t}, \quad \text{with complex scalars } d_1,\, d_2.$$

Fundamental set of *real* eigenfunctions is: $\mathsf{Re}[\mathbf{v}e^{\lambda t}] = \mathsf{Re}[\overline{\mathbf{v}}e^{\overline{\lambda} t}]$, $\mathsf{Im}[\mathbf{v}e^{\lambda t}] = -\mathsf{Im}[\overline{\mathbf{v}}e^{\overline{\lambda} t}]$.

A general *real* solution is then

$$\mathbf{x}(t) = c_1\,\mathsf{Re}[\mathbf{v}e^{\lambda t}] + c_2\,\mathsf{Im}[\mathbf{v}e^{\lambda t}], \quad \text{with real scalars } c_1,\, c_2.$$

The origin $\mathbf{0}$ is a spiral point. If $\lambda = a + bi$ and $a > 0$ then the trajectories spiral outward by a factor e^{at}, if $a < 0$ then the trajectories spiral inward. If $\lambda = ib$ $(b \neq 0)$, then the trajectories rotate around the origin.

Appendix NUMERICAL TOPICS

Optional (the appendix doesn't belong to the course *wi1 277lr*)

When using algorithms for large-scale problems one has to consider questions of computational efficiency and propagation effects of round-off errors in the computations. These algorithms are therefore different from the algorithms used for explaining concepts of linear algebra and making simple computations by hand.

- In solving a system of linear equations the strategy of partial pivoting is used to reduce roundoff errors in calculations.
- In practical numerical work A^{-1} is seldom computed, since solving $A\mathbf{x} = \mathbf{b}$ by row reduction costs less arithmetic operations and may be more accurate.
- The larger the condition number of a square matrix, the closer the matrix is to being singular (non-invertible). Matrix computations with nearly singular or ill-conditioned matrices can produce substantial error.
- Methods of matrix factorization (expression of A as product of matrices) are important for fast numerical computations.

LU Factorization

$A = LU$ with L a lower triangular matrix with 1's on the diagonal, U an echelon form of A (assume no row interchanges needed). Solving $A\mathbf{x} = \mathbf{b}$ is equivalent to solving

$$L\mathbf{y} = \mathbf{b}\,,\ U\mathbf{x} = \mathbf{y}$$

(when A is sparse, this solving is much faster than using A^{-1}).

Algorithm for constructing the LU factorization:
1. Reduce A to an echelon form U by a sequence of row replacement operations
2. Place entries in L such that this sequence of row operations reduces L to I.

QR Factorization

Given: A has linearly independent columns. Then $A = QR$ where the columns of Q form an orthonormal basis for $\mathrm{Col}A$ and R is an upper triangular invertible matrix with positive entries on its diagonal. Remark: $R = Q^T A$.

When the orthonormal basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ is constructed by the Gram-Schmidt process then $Q = [\mathbf{u}_1\, \mathbf{u}_2 \cdots \mathbf{u}_n]$ (with an appropriate sign for each column).

The equation $A\mathbf{x} = \mathbf{b}$ has a unique least-squares solution which can be computed by solving exactly $R\mathbf{x} = Q^T\mathbf{b}$.

There is a QR algorithm for estimating eigenvalues:

$$A = Q_1R_1 \Rightarrow A_1 = R_1Q_1 = Q_2R_2 \Rightarrow A_2 = R_2Q_2 = Q_3R_3 \Rightarrow \cdots$$

A is similar to $A_1, A_2, \ldots$ and A_k becomes almost upper triangular with diagonal entries that approach the eigenvalues of A.

- Iterative estimates for eigenvalues

  Power method for estimating a strictly dominant eigenvalue
  Assume A is diagonalizable, with a basis of eigenvectors $\mathbf{v}_1, \cdots, \mathbf{v}_n$ and corresponding eigenvalues $\lambda_1, \cdots$ where $|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|$.

  When $\mathbf{x}_0 = c_1\mathbf{v}_1 + \cdots$ then $\frac{1}{(\lambda_1)^k} A^k\mathbf{x}_0 \to c_1\mathbf{v}_1$ as $k \to \infty$, therefore $\mathbf{x}_k = A^k\mathbf{x}_0$ points almost in the direction of $\mathbf{v}_1$.
  Scale each $\mathbf{x}_k$ to make its largest entry 1. Then the sequence $\{\mathbf{x}_k\}$ will converge to a multiple of $\mathbf{v}_1$ and the largest entry in $A\mathbf{x}_k$ is close to λ_1.

  1. Select $\mathbf{x}_0$ whose largest entry is 1.
  2. For $k = 0, 1, \ldots$ compute: $\mathbf{y}_k = A\mathbf{x}_k$ and then $\mathbf{x}_{k+1} = \frac{1}{\mu_k}\mathbf{y}_k$ (μ_k is entry of $\mathbf{y}_k$ whose absolute value is largest).
  3. Then $\mathbf{x}_k \to c\mathbf{v}_1$ and $\mu_k \to \lambda_1$.

  The inverse method for estimating an eigenvalue λ of A
  Suppose a good initial estimate α of eigenvalue λ is known. Take $B = (A - \alpha I)^{-1}$ and apply the power method to B. The eigenvalues of B are $\frac{1}{\lambda_1 - \alpha}, \cdots, \frac{1}{\lambda_n - \alpha}$.

  0. Select an initial estimate α close to λ.
  1. Select $\mathbf{x}_0$ whose largest entry is 1.
  2. For $k = 0, 1, \ldots$ compute: $\mathbf{y}_k = B\mathbf{x}_k$ that is solve $(A - \alpha I)\mathbf{y}_k = \mathbf{x}_k$.
  Define $\mathbf{x}_{k+1} = \frac{1}{\mu_k}\mathbf{y}_k$. Compute ν_k with $\frac{1}{\mu_k} = \nu_k - \alpha$.
  3. Then $\mathbf{x}_k \to c\mathbf{v}$ and $\nu_k \to \lambda$.

- Singular Value Decomposition
  For any $m \times n$ matrix A a factorization $A = QDP^{-1}$ is possible. A special factorization of this type is the *singular value decomposition.*

  The singular values of A are the square roots of the eigenvalues of A^TA, that is $\sigma_i = \|A\mathbf{v}_i\|$ with $\mathbf{v}_i$ an unit eigenvector of the symmetric A^TA.

The Singular Value Decomposition
Given: the $m \times n$ matrix A with rank r. Then there exists an $m \times n$ matrix Σ with the first r diagonal entries being the nonzero r singular values and further zeros, and an $m \times m$ orthogonal matrix U and an $n \times n$ orthogonal matrix V such that

$$A = U\Sigma V^T$$

Construction of U and V
Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of $A^T A$, and eigenvalues $\lambda_i \neq 0$ for $1 \leq i \leq r$. $\{A\mathbf{v}_1, \ldots, A\mathbf{v}_r\}$ is a basis of $\mathsf{Col}A$.
Normalize each $\mathbf{u}_i := A\mathbf{v}_i$ for all $1 \leq i \leq r$. Extend to an orthonormal basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ of $\mathbb{R}^m$. Take $U = [\mathbf{u}_1 \ \ \mathbf{u}_2 \ \ \ldots \ \ \mathbf{u}_m]$ and $V = [\mathbf{v}_1 \ \ \mathbf{v}_2 \ \ \ldots \ \ \mathbf{v}_n]$.
The columns of U are left singular vectors, the columns of V are right singular vectors of A.