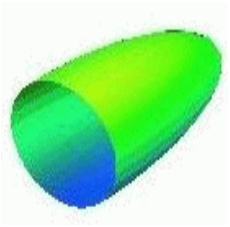
8.5 Eigenanalysis

The subject of **eigenanalysis** seeks to find a coordinate system, in which the solution to an applied problem has a simple expression. Therefore, eigenanalysis might be called *the method of simplifying coordinates*.

Success stories for eigenanalysis include geometric problems, systems of differential equations representing mechanical systems, chemical kinetics or electrical networks, and heat and wave partial differential equations.

Eigenanalysis and footballs

An ellipsoid or *football* is a geometric object described by its **semi-axes** (see Figure 16). In the vector representation, the **semi-axis directions** are unit vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ and the **semi-axis lengths** are the constants a, b, c. The vectors $a\mathbf{v}_1$, $b\mathbf{v}_2$, $c\mathbf{v}_3$ form an **orthogonal triad**.



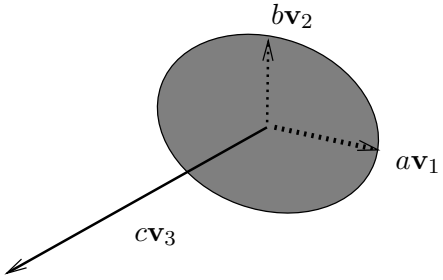


Figure 16. An ellipsoid is built from orthonormal semi-axis directions $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ and the semi-axis lengths a, b, c.

Two vectors $\mathbf{a}$, $\mathbf{b}$ are *orthogonal* if both are nonzero and their dot product $\mathbf{a} \cdot \mathbf{b}$ is zero. Vectors are **orthonormal** if each has norm one and they are pairwise orthogonal. The orthogonal triad is an **invariant** of the ellipsoid's algebraic representations. Algebra does not change the triad: the invariants $a\mathbf{v}_1$, $b\mathbf{v}_2$, $c\mathbf{v}_3$ must somehow be **hidden** in the equations that represent the football.

Algebraic eigenanalysis finds the hidden invariant triad $a\mathbf{v}_1$, $b\mathbf{v}_2$, $c\mathbf{v}_3$ from the ellipsoid's algebraic equations. Suppose, for instance, that the equation of the ellipsoid is supplied as

$$x^2 + 4y^2 + xy + 4z^2 = 16.$$

A symmetric matrix A is constructed in order to write the equation in the form $\mathbf{X}^T A \mathbf{X} = 16$, where $\mathbf{X}$ has components x, y, z. The replacement

equation is[1]

$$(1) \qquad \begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} 1 & 1/2 & 0 \\ 1/2 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 16.$$

It is the 3×3 symmetric matrix A in (1) that is subjected to algebraic eigenanalysis. The matrix calculation will compute the unit directions $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, called the **hidden vectors** or **eigenvectors**. The semi-axis lengths a, b, c are computed at the same time, by finding the **hidden values** or **eigenvalues** λ_1, λ_2, λ_3, known to satisfy the relations

$$\lambda_1 = \frac{16}{a^2}, \quad \lambda_2 = \frac{16}{b^2}, \quad \lambda_3 = \frac{16}{c^2}\,.$$

For the illustration, the football dimensions are $a = 2$, $b = 1.98$, $c = 4.17$. Details of the computation are delayed until page 279.

Coupled and uncoupled systems

The linear system of differential equations

$$(2) \qquad \begin{aligned} x_1' &= -x_1 - x_3, \\ x_2' &= 4x_1 - x_2 - 3x_3, \\ x_3' &= 2x_1 - 4x_3, \end{aligned}$$

is called **coupled**, whereas the linear system of growth-decay equations

$$(3) \qquad \begin{aligned} y_1' &= -3y_1, \\ y_2' &= -y_2, \\ y_3' &= -2y_3, \end{aligned}$$

is called **uncoupled**. The terminology *uncoupled* means that each differential equation in system (3) depends on *exactly one variable*, e.g., $y_1' = -3y_1$ depends only on variable y_1. In a *coupled* system, one of the equations must involve two or more variables.

Matrix characterization of coupled and uncoupled systems. Coupled system (2) and uncoupled system (3) can be written in matrix form, $\mathbf{x}' = A\mathbf{x}$ and $\mathbf{y}' = D\mathbf{y}$, with coefficient matrices

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

If the coefficient matrix is **diagonal**, then the system is uncoupled. If the coefficient matrix is **not diagonal**, then one of the corresponding differential equations involves two or more variables and the system is called coupled or **cross-coupled**.

[1]The reader should pause here and multiply matrices in order to verify this statement. Halving of the entries corresponding to cross-terms generalizes to any ellipsoid.

Solving uncoupled systems

An uncoupled system consists of independent growth-decay equations of the form $u' = au$. The recipe solution $u = ce^{at}$ then leads to the general solution of the system of equations. For instance, system (3) has general solution

$$\begin{aligned} y_1 &= c_1 e^{-3t}, \\ y_2 &= c_2 e^{-t}, \\ y_3 &= c_3 e^{-2t}, \end{aligned} \tag{4}$$

where c_1, c_2, c_3 are **arbitrary constants**. The number of constants equals the dimension of the diagonal matrix D.

Change of coordinates

A change coordinates from $\mathbf{y}$ to $\mathbf{x}$ is a linear algebraic equation $\mathbf{x} = P\mathbf{y}$ where the $n \times n$ matrix P is required to be invertible ($\det(P) \neq 0$). To illustrate, a change of coordinates from $\mathbf{y}$ to $\mathbf{x}$ is given by the linear equations

$$\mathbf{x} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & -1 \\ 2 & 0 & 1 \end{pmatrix} \mathbf{y} \quad \text{or} \quad \begin{cases} x_1 &= y_1 + y_3, \\ x_2 &= y_1 + y_2 - y_3, \\ x_3 &= 2y_1 + y_3. \end{cases} \tag{5}$$

An illustration

A general method exists to construct rich examples of cross-coupled systems. The idea is to substitute a change of variables into a given uncoupled system. To illustrate, substitute into uncoupled system (3) the change of variable equations (5). Then[2]

$$\mathbf{x}' = \begin{pmatrix} -1 & 0 & -1 \\ 4 & -1 & -3 \\ 2 & 0 & -4 \end{pmatrix} \mathbf{x} \quad \text{or} \quad \begin{cases} x_1' = -x_1 - x_3, \\ x_2' = 4x_1 - x_2 - 3x_3, \\ x_3' = 2x_1 - 4x_3. \end{cases} \tag{6}$$

This **cross-coupled** system can be solved using relations (5), (4) and $\mathbf{x} = P\mathbf{y}$ to give the general solution

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & -1 \\ 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} c_1 e^{-3t} \\ c_2 e^{-t} \\ c_3 e^{-2t} \end{pmatrix}. \tag{7}$$

[2]Can't figure out how to compute the coefficients? Use equation (8).

Changing uncoupled systems to cross-coupled systems. Consider a diagonal system $\mathbf{y}' = D\mathbf{y}$, like (3), and a change of variables $\mathbf{x} = P\mathbf{y}$, like (5). Differential calculus applies to give

$$\begin{aligned} \mathbf{x}' &= (P\mathbf{y})' \\ &= P\mathbf{y}' \\ &= PD\mathbf{y} \\ &= PDP^{-1}\mathbf{x}. \end{aligned} \tag{8}$$

The matrix $A = PDP^{-1}$ is *not diagonal* in general, and therefore the change of variables produces a cross-coupled system.

Changing cross-coupled systems to uncoupled systems. We ask this question, motivated by the above calculations:

> Can every cross-coupled system be subjected to a change of variables which converts the system into a completely uncoupled system?

Under certain circumstances, this is true, and it leads to an elegant and especially simple expression for the general solution of the differential system, as in (7).

Eigenanalysis provides a solution to the change of variable question, which distills the infinity of different cross-coupled systems $\mathbf{x}' = A\mathbf{x}$ into a precious few, **classified** by a corresponding uncoupled system $\mathbf{y}' = D\mathbf{y}$, where $A = PDP^{-1}$. The **task of eigenanalysis** is to simultaneously calculate from a cross-coupled system $\mathbf{x}' = A\mathbf{x}$ the change of variables $\mathbf{x} = P\mathbf{y}$ and the diagonal matrix D in the uncoupled system $\mathbf{y}' = D\mathbf{y}$

The **eigenanalysis coordinate system** is the set of n independent vectors extracted from the columns of P. In this coordinate system, the cross-coupled differential system (2) simplifies into a system of uncoupled growth-decay equations (3). Hence the terminology, *the method of simplifying coordinates*.

The algebraic eigenanalysis problem

Given a square matrix A, matrix eigenanalysis seeks to decompose A as a product $A = PDP^{-1}$ with P invertible and D diagonal.

The decomposition is motivated by the preceding discussion of the cross-coupled system $\mathbf{x}' = A\mathbf{x}$ and the diagonal system $\mathbf{y}' = D\mathbf{y}$. We think of A as the cross-coupled system to be solved and the diagonal matrix D as the simplified uncoupled system.

Decomposition $A = PDP^{-1}$. The equation $A = PDP^{-1}$ is equivalent to $AP = PD$. Let $D = \textbf{diag}(\lambda_1, \ldots, \lambda_n)$ and denote the columns of P by $\mathbf{P}_1, \ldots, \mathbf{P}_n$. The columns of AP are the vectors $A\mathbf{P}_1, \ldots, A\mathbf{P}_n$. The columns of PD are $\lambda_1\mathbf{P}_1, \ldots, \lambda_n\mathbf{P}_n$. Therefore, the decomposition $A = PAP^{-1}$ is found from the system of vector equations

$$
(9) \qquad \begin{aligned} A\mathbf{P}_1 &= \lambda_1\mathbf{P}_1, \\ A\mathbf{P}_2 &= \lambda_2\mathbf{P}_2, \\ &\vdots \\ A\mathbf{P}_n &= \lambda_n\mathbf{P}_n. \end{aligned}
$$

Algebraic eigenanalysis is the **atomic** version of (9), isolated as follows.

Definition 5 (Eigenpair)
A pair $(\mathbf{x}, \lambda)$, where $\mathbf{x}$ is a vector and λ is a complex number, is called an **eigenpair** of the $n \times n$ matrix A provided

$$
(10) \qquad A\mathbf{x} = \lambda\mathbf{x} \quad (\mathbf{x} \neq \mathbf{0} \text{ required}).
$$

The **nonzero** requirement is the result of seeking directions for a coordinate system: the zero vector is not a direction. Any vector $\mathbf{x} \neq \mathbf{0}$ that satisfies (10) is called an **eigenvector** for λ and the value λ is called an **eigenvalue** of the square matrix A.

The decomposition $A = PDP^{-1}$, which is equivalent to solving system (9), requires finding all eigenpairs of the $n \times n$ matrix A. If A has n eigenpairs and n independent eigenvectors, then (9) is solved by constructing D as the diagonal matrix of eigenvalues and P as the matrix of corresponding eigenvectors.

The matrix eigenanalysis method

Eigenpairs $(\mathbf{x}, \lambda)$ of a square matrix A are found by the following fundamental result, which reduces the calculation to college algebra and the **rref**-method.

Theorem 9 (Fundamental eigenanalysis theorem)
An eigenpair $(\mathbf{x}, \lambda)$ of a square matrix A is found by the following two-step algorithm:

> **Step 1 (college algebra)**. Solve for eigenvalues λ in the nth order polynomial equation $\det(A - \lambda I) = 0$.
>
> **Step 2 (linear algebra)**. For a given root λ, a corresponding eigenvector $\mathbf{x} \neq \mathbf{0}$ is found by applying the **rref** method to the set of homogeneous linear equations $(A - \lambda I)\mathbf{x} = \mathbf{0}$.

Proof: The equation $A\mathbf{x} = \lambda\mathbf{x}$ is equivalent to $(A - \lambda I)\mathbf{x} = \mathbf{0}$. The latter is a set of homogeneous equations, which always has a solution, that is, consistency is never an issue.

Fix λ and define $B = A - \lambda I$. We show that an eigenpair $(\lambda, \mathbf{x})$ exists with $\mathbf{x} \neq \mathbf{0}$ if and only if $\det(B) = 0$, i.e., $\det(A-\lambda I) = 0$. There is a unique solution $\mathbf{x}$ to the homogeneous equation $B\mathbf{x} = \mathbf{0}$ exactly when Cramer's rule applies, in which case $\mathbf{x} = \mathbf{0}$ is the unique solution. All that Cramer's rule requires is $\det(B) \neq 0$. Therefore, an eigenpair exists exactly when Cramer's rule fails to apply, which is when the determinant of coefficients is zero: $\det(B) = 0$.

A basis of eigenvectors for λ is derived from the parametric solution to the system of equations $(A - \lambda I)\mathbf{x} = \mathbf{0}$. The **rref** method produces systematically a reduced echelon system from which the parametric solution $\mathbf{x}$ is written, depending on parameters $t_1, \ldots, t_k$. Since there is never a unique solution, at least one parameter is required. The *basis* of eigenvectors is obtained from the parametric solution (e.g., $\partial_{t_1}\mathbf{x}, \ldots, \partial_{t_k}\mathbf{x}$).

Orthogonal triad computation. Let's compute the semiaxis directions $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ for the ellipsoid $x^2 + 4y^2 + xy + 4z^2 = 16$. To be applied is Theorem 9. As explained earlier, the starting point is the symmetric matrix

$$A = \begin{pmatrix} 1 & 1/2 & 0 \\ 1/2 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix},$$

which represents the ellipsoid equation as a quadratic form $X^T A X = 16$.

College algebra. The **characteristic polynomial** $\det(A - \lambda I) = 0$ determines the eigenvalues or hidden values of the matrix A. By cofactor expansion, this polynomial equation is

$$(4 - \lambda)((1 - \lambda)(4 - \lambda) - 1/4) = 0$$

with roots 4, $5/2 + \sqrt{10}/2$, $5/2 - \sqrt{10}/2$.

RREF method. It will be shown that the eigenpairs are

$$\lambda_1 = 4, \quad \mathbf{x}_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

$$\lambda_2 = \frac{5 + \sqrt{10}}{2}, \quad \mathbf{x}_2 = \begin{pmatrix} \sqrt{10} - 3 \\ 1 \\ 0 \end{pmatrix},$$

$$\lambda_3 = \frac{5 - \sqrt{10}}{2}, \quad \mathbf{x}_3 = \begin{pmatrix} \sqrt{10} + 3 \\ -1 \\ 0 \end{pmatrix}.$$

The vector norms of the eigenvectors are given by $\|\mathbf{x}_1\| = 1$, $\|\mathbf{x}_2\| = \sqrt{20 + 6\sqrt{10}}$, $\|\mathbf{x}_3\| = \sqrt{20 - 6\sqrt{10}}$. The orthonormal semi-axis direc-

tions $\mathbf{v}_k = \mathbf{x}_k/\|\mathbf{x}_k\|$, $k = 1, 2, 3$, are then given by the formulas

$$\mathbf{v}_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} \frac{\sqrt{10}-3}{\sqrt{20-6\sqrt{10}}} \\ \frac{1}{\sqrt{20-6\sqrt{10}}} \\ 0 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} \frac{\sqrt{10}+3}{\sqrt{20+6\sqrt{10}}} \\ \frac{-1}{\sqrt{20+6\sqrt{10}}} \\ 0 \end{pmatrix}.$$

The details of the **rref** method:

$$\mathbf{aug}(A - \lambda_1 I, \mathbf{0}) = \left(\begin{array}{ccc|c} 1-4 & 1/2 & 0 & 0 \\ 1/2 & 4-4 & 0 & 0 \\ 0 & 0 & 4-4 & 0 \end{array}\right)$$

$$\approx \left(\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \quad \text{Combo, multiply and swap rules. Found } \textbf{rref}.$$

$$\mathbf{aug}(A - \lambda_2 I, \mathbf{0}) = \left(\begin{array}{ccc|c} \frac{-3-\sqrt{10}}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{3-\sqrt{10}}{2} & 0 & 0 \\ 0 & 0 & \frac{3-\sqrt{10}}{2} & 0 \end{array}\right)$$

$$\approx \left(\begin{array}{ccc|c} 1 & 3-\sqrt{10} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \quad \text{Mult, combo and swap rules.}$$

$$\mathbf{aug}(A - \lambda_3 I, \mathbf{0}) = \left(\begin{array}{ccc|c} \frac{-3+\sqrt{10}}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{3+\sqrt{10}}{2} & 0 & 0 \\ 0 & 0 & \frac{3+\sqrt{10}}{2} & 0 \end{array}\right)$$

$$\approx \left(\begin{array}{ccc|c} 1 & 3+\sqrt{10} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \quad \text{Mult, combo and swap rules.}$$

Solving the corresponding reduced echelon systems gives the preceding formulas for the eigenvectors $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$. The equation for the ellipsoid is $\lambda_1 X^2 + \lambda_2 Y^2 + \lambda_3 Z^2 = 16$, where the multipliers of the square terms are the eigenvalues of A and X, Y, Z make up the new coordinates determined by the eigenvectors of A. This equation is easily re-written in the form $X^2/a^2 + Y^2/b^2 + Z^2/c^2 = 1$, provided the semi-axis lengths a, b, c are defined by the relations $a^2 = 16/\lambda_1$, $b^2 = 16/\lambda_2$, $c^2 = 16/\lambda_3$. It is a short computation to obtain $a = 2$, $b = 1.98$, $c = 4.17$.

The ellipse and eigenanalysis

An ellipse equation in **standard form** is $\lambda_1 x^2 + \lambda_2 y^2 = 1$, where $\lambda_1 = 1/a^2$, $\lambda_2 = 1/b^2$ are expressed in terms of the semi-axis lengths a, b. The

expression $\lambda_1 x^2 + \lambda_2 y^2$ is called a **quadratic form**. The study of the ellipse $\lambda_1 x^2 + \lambda_2 y^2 = 1$ is equivalent to the study of the quadratic form equation

$$\mathbf{r}^T D\mathbf{r} = 1, \quad \text{where} \quad \mathbf{r} = \begin{pmatrix} x \\ y \end{pmatrix}, \quad D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

Cross-terms. An ellipse equation may be represented in a uv-coordinate system in a form which has a cross-term uv, e.g., $4u^2 + 8uv + 10v^2 = 5$. The expression $4u^2 + 8uv + 10v^2$ is again called a quadratic form. Calculus courses provide methods to eliminate the cross-term and represent the equation in standard form, by a **rotation**

$$\begin{pmatrix} u \\ v \end{pmatrix} = R \begin{pmatrix} x \\ y \end{pmatrix}, \quad R = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.$$

The angle θ in the rotation matrix R represents the counterclockwise rotation of uv-coordinates into standard xy-coordinates.

Eigenanalysis computes the rotation angle θ through the columns of R, which are the unit semi-axis directions of the unrotated ellipse. If the quadratic form $4u^2 + 8uv + 10v^2$ is represented as $\mathbf{r}^T A\,\mathbf{r}$, then

$$\mathbf{r} = \begin{pmatrix} u \\ v \end{pmatrix}, \quad A = \begin{pmatrix} 4 & 4 \\ 4 & 10 \end{pmatrix}.$$

An eigenanalysis of A will give the orthogonal unit semi-axis directions $\mathbf{v}_1$, $\mathbf{v}_2$ as eigenvectors of A, which are the columns of the rotation matrix R. The components of an eigenvector can be used to determine angle θ.

The semi-axis lengths $a \approx 1.55$, $b \approx 0.63$ of the ellipse $4u^2+8uv+10v^2 = 5$ are computed from the eigenvalues $\lambda_1 = 12$, $\lambda_2 = 2$ of matrix A by the equations

$$\frac{\lambda_1}{5} = \frac{1}{a^2}, \quad \frac{\lambda_2}{5} = \frac{1}{b^2}.$$

Geometrically, the ellipse $4u^2 + 8uv + 10v^2 = 5$ is completely determined by the orthogonal semi-axis vectors $a\mathbf{v}_1$, $b\mathbf{v}_2$. The rotation R is a rigid motion which maps these vectors into $a\mathbf{i}$, $b\mathbf{j}$, where $\mathbf{i}$ and $\mathbf{j}$ are the standard unit vectors in the plane.

The θ-rotation R maps $4u^2+8uv+10v^2 = 5$ into the xy-equation $\lambda_1 x^2 + \lambda_2 y^2 = 5$, where λ_1, λ_2 are the eigenvalues of A. To see why, let $\mathbf{r} = R\mathbf{s}$ where $\mathbf{s} = \begin{pmatrix} x & y \end{pmatrix}^T$. Then $\mathbf{r}^T A\mathbf{r} = \mathbf{s}^T(R^T AR)\mathbf{s}$. Using $R^T = R^{-1}$ (equivalently, $R^T R = I$) gives $R^T AR = \mathbf{diag}(\lambda_1, \lambda_2)$ and finally $\mathbf{r}^T A\mathbf{r} = \lambda_1 x^2 + \lambda_2 y^2$.

Diagonalization

A system of differential equations $\mathbf{x}' = A\mathbf{x}$ can be transformed to an uncoupled system $\mathbf{y}' = \mathbf{diag}(\lambda_1, \ldots, \lambda_n)\mathbf{y}$ by a change of variables $\mathbf{x} = P\mathbf{y}$, provided P is invertible and A satisfies the relation

$$AP = P\,\mathbf{diag}(\lambda_1, \ldots, \lambda_n). \tag{11}$$

A matrix A is said to be **diagonalizable** provided (11) holds. This equation is equivalent to the system of equations $A\mathbf{P}_k = \lambda_k \mathbf{P}_k$, $k = 1, \ldots, n$, where $\mathbf{P}_1, \ldots, \mathbf{P}_n$ are the columns of matrix P. Since P is assumed invertible, each of its columns are nonzero, and therefore $(\lambda_k, \mathbf{P}_k)$ is an eigenpair of A, $1 \le k \le n$. The values λ_k need not be distinct (e.g., all $\lambda_k = 1$ if A is the identity). This proves:

Theorem 10 (Diagonalization)
An $n \times n$ matrix A is diagonalizable if and only if A has n eigenpairs $(\lambda_k, \mathbf{x}_k)$, $1 \le k \le n$, with $\mathbf{x}_1, \ldots, \mathbf{x}_n$ independent.

An eigenpair $(\lambda, \mathbf{x})$ of A can always be selected so that $\|\mathbf{x}\| = 1$. For example, replace $\mathbf{x}$ by $c\mathbf{x}$ where $c = 1/\|\mathbf{x}\|$. By this small change, it can be assumed that P has columns of unit length.

Theorem 11 (Orthogonality of eigenvectors)
Assume that $n \times n$ matrix A is **symmetric**, $A^T = A$. If $(\alpha, \mathbf{x})$ and $(\beta, \mathbf{y})$ are eigenpairs of A with $\alpha \neq \beta$, then $\mathbf{x} \cdot \mathbf{y} = 0$, that is, $\mathbf{x}$ and $\mathbf{y}$ are orthogonal.

Proof: To prove this result, compute $\alpha \mathbf{x} \cdot \mathbf{y} = (A\mathbf{x})^T \mathbf{y} = \mathbf{x}^T A^T \mathbf{y} = \mathbf{x}^T A \mathbf{y}$. Similarly, $\beta \mathbf{x} \cdot \mathbf{y} = \mathbf{x}^T A \mathbf{y}$. Together, these relations imply $(\alpha - \beta)\mathbf{x} \cdot \mathbf{y} = 0$, giving $\mathbf{x} \cdot \mathbf{y} = 0$ due to $\alpha \neq \beta$.

The Gram-Schmidt process. The eigenvectors of a symmetric matrix A may be constructed to be orthogonal. First of all, observe that eigenvectors corresponding to distinct eigenvalues are orthogonal by Theorem 11. It remains to construct from k independent eigenvectors $\mathbf{x}_1, \ldots, \mathbf{x}_k$, corresponding to a single eigenvalue λ, another set of independent eigenvectors $\mathbf{y}_1, \ldots, \mathbf{y}_k$ for λ which are pairwise orthogonal. The idea, due to Gram-Schmidt, applies to any set of k independent vectors.

Application of the Gram-Schmidt process can be illustrated by example: let $(-1, \mathbf{v}_1)$, $(2, \mathbf{v}_2)$, $(2, \mathbf{v}_3)$, $(2, \mathbf{v}_4)$ be eigenpairs of a 4×4 symmetric matrix A. Then $\mathbf{v}_1$ is orthogonal to $\mathbf{v}_2$, $\mathbf{v}_3$, $\mathbf{v}_4$. The vectors $\mathbf{v}_2$, $\mathbf{v}_3$, $\mathbf{v}_4$ belong to eigenvalue $\lambda = 2$, but they are not necessarily orthogonal. The Gram-Schmidt process replaces these vectors by $\mathbf{y}_2$, $\mathbf{y}_3$, $\mathbf{y}_4$ which are pairwise orthogonal. The result is that eigenvectors $\mathbf{v}_1$, $\mathbf{y}_2$, $\mathbf{y}_3$, $\mathbf{y}_4$ are pairwise orthogonal.

Theorem 12 (Gram-Schmidt)
Let $\mathbf{x}_1, \ldots, \mathbf{x}_k$ be independent n-vectors. The set of vectors $\mathbf{y}_1, \ldots, \mathbf{y}_k$ constructed below as linear combinations of $\mathbf{x}_1, \ldots, \mathbf{x}_k$ are pairwise orthogonal and independent.

$$\begin{aligned}
\mathbf{y}_1 &= \mathbf{x}_1 \\
\mathbf{y}_2 &= \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{y}_1}{\mathbf{y}_1 \cdot \mathbf{y}_1}\mathbf{y}_1 \\
\mathbf{y}_3 &= \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{y}_1}{\mathbf{y}_1 \cdot \mathbf{y}_1}\mathbf{y}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{y}_2}{\mathbf{y}_2 \cdot \mathbf{y}_2}\mathbf{y}_2 \\
&\vdots \\
\mathbf{y}_k &= \mathbf{x}_k - \frac{\mathbf{x}_k \cdot \mathbf{y}_1}{\mathbf{y}_1 \cdot \mathbf{y}_1}\mathbf{y}_1 - \cdots - \frac{\mathbf{x}_k \cdot \mathbf{y}_{k-1}}{\mathbf{y}_{k-1} \cdot \mathbf{y}_{k-1}}\mathbf{y}_{k-1}
\end{aligned}$$

Proof: Let's begin with a lemma: *Any set of nonzero orthogonal vectors* $\mathbf{y}_1, \ldots, \mathbf{y}_k$ *are independent.* Assume the relation $c_1\mathbf{y}_1 + \cdots + c_k\mathbf{y}_k = \mathbf{0}$. Take the dot product of this relation with $\mathbf{y}_j$. By orthogonality, $c_j\, \mathbf{y}_j \cdot \mathbf{y}_j = 0$, and since $\mathbf{y}_j \neq \mathbf{0}$, cancellation gives $c_j = 0$ for $1 \leq j \leq k$. Hence $\mathbf{y}_1, \ldots, \mathbf{y}_k$ are independent.

Induction will be applied on k to show that $\mathbf{y}_1, \ldots, \mathbf{y}_k$ are nonzero and orthogonal. If $k = 1$, then there is just one nonzero vector constructed $\mathbf{y}_1 = \mathbf{x}_1$. Orthogonality for $k = 1$ is not discussed because there are no pairs to test. Assume the result holds for $k-1$ vectors. Let's verify that it holds for k vectors, $k > 1$. Assume orthogonality $\mathbf{y}_i \cdot \mathbf{y}_j = 0$ and $\mathbf{y}_i \neq \mathbf{0}$ for $1 \leq i, j \leq k-1$. It remains to test $\mathbf{y}_i \cdot \mathbf{y}_k = 0$ for $1 \leq i \leq k-1$ and $\mathbf{y}_k \neq \mathbf{0}$. The test depends upon the identity

$$\mathbf{y}_i \cdot \mathbf{y}_k = \mathbf{y}_i \cdot \mathbf{x}_k - \sum_{j=1}^{k-1} \frac{\mathbf{x}_k \cdot \mathbf{y}_j}{\mathbf{y}_j \cdot \mathbf{y}_j}\, \mathbf{y}_i \cdot \mathbf{y}_j,$$

which is obtained from the formula for $\mathbf{y}_k$ by taking the dot product with $\mathbf{y}_i$. In the identity, $\mathbf{y}_i \cdot \mathbf{y}_j = 0$ by the induction hypothesis for $1 \leq j \leq k-1$ and $j \neq i$. Therefore, the summation in the identity contains just the term for index $j = i$, and the contribution is $\mathbf{y}_i \cdot \mathbf{x}_k$. This contribution cancels the leading term on the right in the identity, resulting in the orthogonality relation $\mathbf{y}_i \cdot \mathbf{y}_k = 0$. If $\mathbf{y}_k = \mathbf{0}$, then $\mathbf{x}_k$ is a linear combination of $\mathbf{y}_1, \ldots, \mathbf{y}_{k-1}$. But each $\mathbf{y}_j$ is a linear combination of $\{\mathbf{x}_i\}_{i=1}^{j}$, therefore $\mathbf{y}_k = \mathbf{0}$ implies $\mathbf{x}_k$ is a linear combination of $\mathbf{x}_1, \ldots, \mathbf{x}_{k-1}$, a contradiction to the independence of $\{\mathbf{x}_i\}_{i=1}^{k}$. The proof is complete.

Cayley-Hamilton identity

A celebrated and deep result for powers of matrices was discovered by Cayley and Hamilton [ref?], which says that an $n \times n$ matrix A satisfies its own characteristic equation. More precisely:

Theorem 13 (Cayley-Hamilton)
Let $\det(A - \lambda I)$ be expanded as the nth degree polynomial

$$p(\lambda) = \sum_{j=0}^{n} c_j \lambda^j,$$

for some coefficients $c_0, \ldots, c_{n-1}$ and $c_n = (-1)^n$. Then A satisfies the equation $p(\lambda) = 0$, that is,

$$p(A) \equiv \sum_{j=0}^{n} c_j A^j = 0.$$

In factored form in terms of the eigenvalues $\{\lambda_j\}_{j=1}^n$ (duplicates possible),

$$(-1)^n (A - \lambda_1 I)(A - \lambda_2 I) \cdots (A - \lambda_n I) = 0.$$

Proof: If A is diagonalizable, $AP = P\,\mathbf{diag}(\lambda_1, \ldots, \lambda_n)$, then the proof is obtained from the simple expansion

$$A^j = P\,\mathbf{diag}(\lambda_1^j, \ldots, \lambda_n^j) P^{-1},$$

because summing across this identity leads to

$$\begin{array}{rcl} p(A) & = & \sum_{j=0}^n c_j A^j \\ & = & P \left(\sum_{j=0}^n c_j\, \mathbf{diag}(\lambda_1^j, \ldots, \lambda_n^j) \right) P^{-1} \\ & = & P\,\mathbf{diag}(p(\lambda_1), \ldots, p(\lambda_n)) P^{-1} \\ & = & P\,\mathbf{diag}(0, \ldots, 0) P^{-1} \\ & = & 0. \end{array}$$

If A is not diagonalizable, then this proof fails. To handle the general case, we apply a deep linear algebra result, known as **Jordan's theorem**, which says that $A = PJP^{-1}$ where J is *upper triangular*, instead of *diagonal*. The not necessarily distinct eigenvalues $\lambda_1, \ldots, \lambda_n$ of A appear on the diagonal of J. Using this result, define

$$A_\epsilon = P(J + \epsilon\,\mathbf{diag}(1, 2, \ldots, n)) P^{-1}.$$

For small $\epsilon > 0$, the matrix A_ϵ has distinct eigenvalues $\lambda_j + \epsilon j$, $1 \le j \le n$. Then the diagonalizable case implies that A_ϵ satisfies its characteristic equation $p_\epsilon(\lambda) = \det(A_\epsilon - \lambda I) = 0$. Use $0 = \lim_{\epsilon \to 0} p_\epsilon(A_\epsilon) = p(A)$ to complete the proof.

Generalized eigenanalysis

The main result of generalized eigenanalysis is the equation

$$A = PJP^{-1},$$

valid for any real or complex square matrix A. The matrix J is an upper triangular matrix called the **Jordan form** of the matrix A and the columns of P are called **generalized eigenvectors** of A.

Due to the triangular form of J, all eigenvalues of A appear on the main diagonal of J, which gives J the generic form

$$J = \begin{pmatrix} \lambda_1 & c_{12} & c_{13} & \cdots & c_{1n} \\ 0 & \lambda_2 & c_{23} & \cdots & c_{2n} \\ 0 & 0 & \lambda_3 & \cdots & c_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

The columns of P are independent — they form a coordinate system. There is for each eigenvalue λ of A at least one column $\mathbf{x}$ of P satisfying $A\mathbf{x} = \lambda\mathbf{x}$. However, there may be other columns of P that *fail to be eigenvectors*, that is, $A\mathbf{x} = \lambda\mathbf{x}$ may be *false* for many columns $\mathbf{x}$ of P.

Solving triangular differential systems. A matrix differential system $\mathbf{x}'(t) = J\mathbf{x}(t)$ with J upper triangular splits into scalar equations which can be solved by elementary methods for first order scalar differential equations. To illustrate, consider the system

$$\begin{aligned} x_1' &= 3x_1 + x_2 + x_3, \\ x_2' &= 3x_2 + x_3, \\ x_3' &= 2x_3. \end{aligned}$$

The techniques that apply are the growth-decay recipe for $u' = ku$ and the factorization method for $u' = ku + p(t)$. Working backwards from the last equation, using back-substitution, gives

$$\begin{aligned} x_3 &= c_3 e^{2t}, \\ x_2 &= c_2 e^{3t} - c_3 e^{2t}, \\ x_1 &= (c_1 + c_2 t) e^{3t}. \end{aligned}$$

The real Jordan form of A. Given a real matrix A, **generalized eigenanalysis** seeks to find a real invertible matrix P such that $AP = PD$, where D is a **real Jordan block diagonal matrix**.

A **real Jordan block** corresponding to a real eigenvalue λ of A is a matrix

$$B = \mathbf{diag}(\lambda, \ldots, \lambda) + N,$$

where

$$N = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

If $\lambda = a + ib$ is a complex eigenvalue, then in B, λ is replaced by a 2×2 real matrix $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$, and in N, 1 is replaced by the 2×2 matrix $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. The matrix N satisfies $N^k = 0$ for some integer k; such matrices are called **nilpotent**.

A Jordan block system $\mathbf{x}' = B\mathbf{x}$ can be solved by elementary first-order scalar methods, using only the real number system.

=========

The change of coordinates $\mathbf{y} = P\mathbf{x}$ then changes $\mathbf{x}' = A\mathbf{x}$ into the Jordan block diagonal system $\mathbf{y}' = D\mathbf{y}$.