AE 525a/ME 525 Lecture #3: Eigenvalue Problem

Let's consider a matrix $\mathsf{A} \in R^{\mathsf{n}\times\mathsf{n}}$ or $\mathsf{A} \in C^{\mathsf{n}\times\mathsf{n}}$ and the problem

$$\mathsf{Ax} = \lambda \mathsf{x} \tag{1}$$

then λ is the eigenvalue and x is corresponding eigenvector of A. Of special interest are symmetric matrices ($\mathsf{A}^\mathsf{T} = \mathsf{A}$) and Hermitian matrices ($\mathsf{A} = \overline{\mathsf{A}}^\mathsf{T}$).

Let's first consider the case of real matrices. Then the problem can be restated as

$$(\mathsf{A} - \lambda \mathsf{I})\mathsf{x} = \mathsf{0} \tag{2}$$

A nontrivial solution for the homogeneous problem exists if and only if

$$\det(\mathsf{A} - \lambda \mathsf{I}) = 0 \tag{3}$$

which is called the characteristic equation of A. The eigenvalues are the roots of an algebraic equation of order n. The eigenvalues $\lambda_\mathsf{i}; i = 1 : n$ need not be distinct.

Corresponding to each λ_k, there exists at least one vector $\mathsf{x}^{(\mathsf{k})}$ such that

$$\mathsf{Ax}^{(\mathsf{k})} = \lambda_\mathsf{k} \mathsf{x}^{(\mathsf{k})} \tag{4}$$

Let λ_1 and λ_2 be two distinct eigenvalues with corresponding eigenvectors u_1, and u_2, respectively. Then

$$\mathsf{Au}_1 = \lambda_1 \mathsf{u}_1 \tag{5}$$

$$\mathsf{Au}_2 = \lambda_1 \mathsf{u}_2 \tag{6}$$

Then the last two equations provide

$$\begin{aligned} \mathsf{u}_1^\mathsf{T} \mathsf{Au}_2 &= \lambda_1 \mathsf{u}_1^\mathsf{T} \mathsf{u}_2 \\ \mathsf{u}_1^\mathsf{T} \mathsf{Au}_2 &= \lambda_2 \mathsf{u}_1^\mathsf{T} \mathsf{u}_2 \end{aligned}$$

and thus

$$(\lambda_1 - \lambda_2)(\mathsf{u}_1, \mathsf{u}_2) = 0$$

Since the eigenvalues are distinct we get that

$$(\mathsf{u}_1, \mathsf{u}_2) = 0$$

Therefore, we have the result that two eigenvectors of a real symmetric matrix, corresponding to different eigenvalues are orthogonal, i.e.,

$$(\mathsf{u}_\mathsf{m}, \mathsf{u}_\mathsf{n}) = 0; \quad for \quad \lambda_\mathsf{m} \neq \lambda_\mathsf{n} \tag{7}$$

Claim 1 For a symmetric real matrix A corresponding eigenvalues are real.

Proof. Assume that the eigenvalue is complex, i.e., $\lambda = \alpha + i\beta$. Then, $\mathsf{Ax} = \lambda\mathsf{x}$ implies

$$\begin{aligned} \mathsf{x}^{\mathsf{T}}\mathsf{A}\overline{\mathsf{x}} &= \lambda\mathsf{x}^{\mathsf{T}}\overline{\mathsf{x}} \\ \mathsf{x}^{\mathsf{T}}\mathsf{A}\overline{\mathsf{x}} &= \overline{\lambda}\overline{\mathsf{x}}^{\mathsf{T}}\mathsf{x} = \overline{\lambda}\mathsf{x}^{\mathsf{T}}\overline{\mathsf{x}} \end{aligned}$$

so we get

$$(\lambda - \overline{\lambda})(\overline{\mathsf{x}}, \mathsf{x}) = 0$$

or

$$\overline{\lambda} = \lambda$$

■

Claim 2 If an eigenvalue, say, λ_1 of a symmetric matrix is a multiple root of multiplicity s, i.e., if the LHS of $\det(\mathsf{A} - \lambda\mathsf{I}) = 0$ possesses the factor $(\lambda - \lambda_1)^{\mathsf{s}}$, then to λ_1 there correspond s linearly independent eigenvectors, any nontrivial combination of which also have the same property.

We postpone the proof for a while.

Remark 3 The last result may not be true for nonsymmetric matrices. For example, matrix

$$\mathsf{A} = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$$

is nonsymmetric and the corresponding characteristic equation is

$$\begin{aligned} \det(\mathsf{A} - \lambda\mathsf{I}) &= \det\begin{bmatrix} 1-\lambda & 1 \\ -1 & -1-\lambda \end{bmatrix} \\ &= -(1-\lambda^2) + 1 = 0 \end{aligned}$$

Therefore, $\lambda = 0; s = 2$. However, for $\lambda = 0$ we get $\mathsf{Ax} = \mathsf{0}$. Since the $rank\,\mathsf{A} = 1$, the only possible solution is of the form

$$\mathsf{x} = C_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

and thus for eigenvalue $\lambda = 0$ there is only one eigenvector.

Claim 4 It is always possible (will be shown later) to choose s linearly independent vectors corresponding to an eigenvalue of multiplicity s in such a way that they are orthogonal to each other, in addition to being (automatically) orthogonal to all other eigenvectors.

If multiple roots of the characteristic equation are counted separately, we obtain exactly n eigenvalues, and from that we can determine a set of n mutually orthogonal eigenvectors. This set of vectors forms a basis in $n-$space.

If the basis is normalized we get an orthonormal set $\mathsf{e}_i; i = 1 : n$ such that $(\mathsf{e}_i, \mathsf{e}_j) = \delta_{ij}$. Thus for any vector v

$$\mathsf{v} = \alpha_k \mathsf{e}_k; k = 1 : n \tag{8}$$

we get

$$(\mathsf{e}_i, \mathsf{v}) = \boldsymbol{\alpha}_k \delta_{ki} = \alpha_i$$

Consider now the inhomogeneous system of equations

$$\mathsf{Ax} - \lambda \mathsf{x} = \mathsf{c} \tag{9}$$

which for $\mathsf{c} = \mathsf{0}$ reduces to $\mathsf{Ax} = \lambda \mathsf{x}$, where $\mathsf{A} \in R^{n \times n}$. If (9) has a solution then that solution can be expressed as a linear combination of the eigenvectors of A. Assume that n orthonormal eigenvectors $\mathsf{e}_i; i = 1 : n$ are known and thus

$$\mathsf{Ae}_i = \lambda_{\underline{i}} \mathsf{e}_{\underline{i}}; i = 1 : n$$

Consequently,

$$\mathsf{x} = \boldsymbol{\alpha}_k \mathsf{e}_k; k = 1 : n \tag{10}$$

where $\alpha_k; k = 1 : n$ are to be determined. By substituting (10) into (9) we get

$$(\lambda_k - \lambda) \alpha_k \mathsf{e}_k = \mathsf{c}$$

which after inner product with e_i provides

$$(\lambda_i - \lambda) \alpha_i = (\mathsf{e}_i, \mathsf{c})$$

and thus

$$\alpha_i = \frac{(\mathsf{e}_i, \mathsf{c})}{\lambda_i - \lambda}; i = 1 : n \tag{11}$$

Therefore, the solution for the inhomogeneous problem can be written as

$$\mathsf{x} = \sum_{k=1}^{\infty} \frac{(\mathsf{e}_k, \mathsf{c})}{\lambda_k - \lambda} \mathsf{e}_k \tag{12}$$

Apparently, a unique solution exists when $\lambda \neq \lambda_k \equiv$an eigenvalue. If $\lambda = \lambda_p$, no solution exists unless $(\mathsf{e}_p, \mathsf{c}) = 0$. In that case equation $(\lambda_p - \lambda)\alpha_p = (\mathsf{e}_p, \mathsf{c}) = 0$, which implies that α_p can be chosen arbitrarily, so that infinitely many solutions exists.

Remark 5 If $\lambda = 0$ then $\mathsf{Ax} - \lambda \mathsf{x} = \mathsf{c}$ reduces to $\mathsf{Ax} = \mathsf{c}$, which we have studied earlier. This equation has unique solution unless $\lambda = 0$ is an eigenvalues of A, i.e., unless the equation $\mathsf{Ax} = \lambda \mathsf{x} = \mathsf{0}$ has a nontrivial solution (note this requires that $\det \mathsf{A} = 0$ or that $rank(\mathsf{A}) < n$). Then no solution exists unless c is orthogonal to the vectors satisfying $\mathsf{Ax} = \mathsf{0}$ (recall for $\lambda = 0$

$$\mathsf{x} = \sum_{k=1}^{\infty} \frac{(\mathsf{e}_k, \mathsf{c})}{\lambda_k} \mathsf{e}_k$$

1 Orthogonalization of Vector Sets

Often we are faced with the following task. Given a set of s linearly independent vectors $\mathsf{u}_i; i = 1 : s$ form a set of s orthonormal vectors.

We choose any original vector, say, $\mathsf{v}_1 = \mathsf{u}_1$ and thus

$$\mathsf{e}_1 = \frac{\mathsf{u}_1}{\ell(\mathsf{u}_1)}$$

Next, we choose u_2 and form a vector

$$\mathsf{v}_2 = \mathsf{u}_2 - c\mathsf{e}_1$$

Since v_2 must be orthogonal to e_1 we must have

$$\begin{aligned} (\mathsf{e}_1, \mathsf{v}_2) &= 0 = (\mathsf{e}_1, \mathsf{u}_2) - c \Rightarrow \\ c &= (\mathsf{e}_1, \mathsf{u}_2) \end{aligned}$$

and thus

$$\begin{aligned} \mathsf{v}_2 &= \mathsf{u}_2 - (\mathsf{e}_1, \mathsf{u}_2)\mathsf{e}_1 \Rightarrow \\ \mathsf{e}_2 &= \frac{\mathsf{v}_2}{\ell(\mathsf{v}_2)} \end{aligned}$$

Similarly,

$$\mathsf{v}_3 = \mathsf{u}_3 - a\mathsf{e}_1 - b\mathsf{e}_2$$

and since v_3 must be orthogonal to both e_1 and e_2 we get

$$\begin{aligned} a &= (\mathsf{e}_1, \mathsf{u}_3) \\ b &= (\mathsf{e}_2, \mathsf{u}_3) \\ \mathsf{v}_3 &= \mathsf{u}_3 - (\mathsf{e}_1, \mathsf{u}_3)\mathsf{e}_1 - (\mathsf{e}_2, \mathsf{u}_3)\mathsf{e}_2 \Rightarrow \\ \mathsf{e}_3 &= \frac{\mathsf{v}_3}{\ell(\mathsf{v}_3)} \end{aligned}$$

In general we have

$$\begin{aligned} \mathsf{v}_s &= \mathsf{u}_s - \sum_{k=1}^{s-1} (\mathsf{e}_k, \mathsf{u}_s)\mathsf{e}_k \\ \mathsf{e}_s &= \frac{\mathsf{v}_s}{\ell(\mathsf{v}_s)} \end{aligned}$$

This approach is known as Gram-Schmidt orthogonalization procedure.

The Gram-Schmidt procedure will fail if and only if at some stage $\mathsf{v}_r = 0 (= \mathsf{u}_r - \sum_{k=1}^{r-1} (\mathsf{e}_k, \mathsf{u}_r)\mathsf{e}_k)$. However, this implies that u_r is a linear combination of $\mathsf{e}_i; i = 1 : r-1$ which is in contradiction to the initial hypothesis that $\mathsf{u}_i; i = 1 : s$ are linearly independent.

Remark 6 For complex vectors $u_i; i = 1 : s$ the same procedure is valid if Hermitian products and lengths are used.

Example 7

$$
\begin{aligned}
u_1 &= \begin{bmatrix} 1 \\ 1 \\ 1 \\ -1 \end{bmatrix}; u_2 = \begin{bmatrix} 2 \\ -1 \\ -1 \\ 1 \end{bmatrix}; u_3 = \begin{bmatrix} -1 \\ 2 \\ 2 \\ 1 \end{bmatrix} \\
\ell(u_1) &= 2; \ell(u_2) = \sqrt{7}; \ell(u_3) = \sqrt{10} \\
e_1 &= \frac{u_1}{\ell(u_1)} = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 1 \\ -1 \end{bmatrix} \\
v_2 &= u_2 - c e_1 \\
c &= (e_1, u_2) = -1/2 \\
v_2 &= \frac{1}{4}\begin{bmatrix} 9 \\ -3 \\ -3 \\ 3 \end{bmatrix}; e_2 = \frac{v_2}{\ell(v_2)} = \frac{\sqrt{3}}{2}\begin{bmatrix} 1 \\ -1/3 \\ -1/3 \\ 1/3 \end{bmatrix} \\
v_3 &= u_3 - a e_1 - b e_2 \\
a &= (e_1, u_3) = 1; b = (e_2, u_3) = -\sqrt{3} \\
v_3 &= \begin{bmatrix} 0 \\ 1 \\ 1 \\ 2 \end{bmatrix}; e_3 = \frac{v_3}{\ell(v_3)} = \frac{\sqrt{6}}{6}\begin{bmatrix} 0 \\ 1 \\ 1 \\ 2 \end{bmatrix}
\end{aligned}
$$

2 Quadratic Forms

Definition 8 A homogeneous expression of second degree of the form

$$
\begin{aligned}
A &= a_{11}x_1^2 + a_{22}x_2^2 + ... + a_{nn}x_n^2 \\
&\quad + 2a_{12}x_1x_2 + ... + 2a_{n-1,n}x_{n-1}x_n \\
&= x^T A x
\end{aligned}
\tag{13}
$$

is called a quadratic form in $x_1, ..., x_n$. We shall assume that $A \in R^{n \times n}$.

Thus in 2D space $A = const$ represents a second degree curve with center at the origin while in $3D$ space $A = const$ denotes a central quadratic surface with center at O.

Remark 9 If we write

$$y_i = \frac{1}{2}\frac{\partial A}{\partial x_i}$$

then

$$\mathsf{A}\mathsf{x} = \mathsf{y}$$

where $\mathsf{A}^{\mathsf{T}} = \mathsf{A}$ has been used.

Often we would like to reduce quadratic form a to the following

$$A = a'_{11}x'_{11} + a'_{22}x'_{2} + ... + a'_{nn}x'_{n} \tag{14}$$

so that all the cross-product terms are eliminated. Let x and x' are related by

$$\mathsf{x} = \mathsf{Q}\mathsf{x}'; \mathsf{Q} \in R^{n\times n} \tag{15}$$

and then

$$A = \mathsf{x}'^{\mathsf{T}}\mathsf{A}'\mathsf{x}' \tag{16}$$

$$\mathsf{A}' = \mathsf{Q}^{\mathsf{T}}\mathsf{A}\mathsf{Q} \tag{17}$$

If A' is diagonal then there will be no cross products. Thus Eqn.(16) represents a canonical form if A' is a diagonal matrix.

Suppose we know the eigenvalues $\lambda_i; i = 1 : n$ of $\mathsf{A} = \mathsf{A}^{\mathsf{T}} \in R^{n\times n}$ as well as the corresponding eigenvectors $\mathsf{e}_i; i = 1 : n$. Let's construct Q in such a way that the eigenvectors $\mathsf{e}_i; i = 1 : n$ are the successive columns of Q

$$\begin{aligned} \mathsf{Q} &= [\mathsf{e}_1, ..., \mathsf{e}_n] \\ &= \begin{bmatrix} e_{11} & e_{21} & \cdots & e_{n1} \\ e_{12} & e_{22} & & e_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ e_{1n} & e_{2n} & \cdots & e_{nn} \end{bmatrix} \end{aligned} \tag{18}$$

Then

$$\begin{aligned} \mathsf{A}\mathsf{Q} &= \mathsf{A}[\mathsf{e}_1, ..., \mathsf{e}_n] = [\lambda_1\mathsf{e}_1, ..., \lambda_n\mathsf{e}_n] \\ &= [\mathsf{e}_1, ..., \mathsf{e}_n]diag(\lambda_1, ..., \lambda_n) \\ &= \mathsf{Q}diag(\lambda_1, ..., \lambda_n) \end{aligned}$$

Since $\mathsf{e}_i; i = 1 : n$ are linearly independent we have that $\det \mathsf{Q} \neq 0$ so the inverse Q^{-1} exists and thus

$$\mathsf{Q}^{-1}\mathsf{A}\mathsf{Q} = diag(\lambda_1, ..., \lambda_n) \tag{19}$$

If the matrix Q is such that $\mathsf{Q}^{-1} = \mathsf{Q}^{\mathsf{T}}$ the quadratic form (16) is in the canonical form. The matrix Q is called a modal matrix of A.

If $\mathsf{e}_i; i = 1 : n$ are orthonormal then Q is orthonormal.

Remark 10 If $\lambda = 0$ is an eigenvalue of multiplicity s, then the quadratic form A has only $n - s$ nonzero terms

$$A = \lambda_1 x_1'^2 + ... + \lambda_{n-s} x_{n-s}'^2 \tag{20}$$

We have shown before that this case arises if and only if the symmetric matrix A is of rank $n - s = r$.

Remark 11 Since $\mathsf{x}' = \mathsf{Q}^{-1}\mathsf{x} = \mathsf{Q}^{\mathsf{T}}x$, we have that $x'_{\mathsf{i}} = (\mathsf{e}_{\mathsf{i}}, \mathsf{x}); \mathsf{i} = 1 : \mathsf{n}$.

If all eigenvalues are distinct, the modal matrix Q is uniquely determined (except for the ordering of the columns and the algebraic sign associated with each column). However, if an eigenvalue is of multiplicity s, the corresponding s orthogonal unit vectors can be chosen in infinitely many ways (as discussed before).

Remark 12 The matrices for which $\mathsf{Q}^{\mathsf{T}} = \mathsf{Q}^{-1}$ are called orthogonal matrices.

Claim 13 A square matrix is an orthogonal matrix if and only if its columns are mutually orthogonal unit vectors.

Example 14 Consider a quadratic form $A = 25x_1^2 + 34x_2^2 + 41x_3^2 - 24x_2x_3$. Therefore,

$$\begin{aligned}
\mathsf{A} &= \begin{bmatrix} 25 & 0 & 0 \\ 0 & 34 & -12 \\ 0 & -12 & 41 \end{bmatrix} \\
\det(\mathsf{A}-\lambda\mathsf{I}) = (25-\boldsymbol{\lambda})[\quad (43-\lambda)(41-\lambda) - 12^2] &= 0 \\
\lambda_{1,2} &= 25 \\
\lambda_3 &= 50
\end{aligned}$$

For $\lambda = 25$ we get from $\mathsf{Ax} - \lambda\mathsf{x} = 0$

$$\begin{aligned}
x_1 &= C_1 \\
x_2 &= C_2 \\
x_3 &= \frac{3}{4}C_2
\end{aligned}$$

and thus

$$\begin{aligned}
\mathsf{x} &= C_1\mathsf{u}_1 + C_2\mathsf{u}_2 \\
\mathsf{u}_1 &= \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}; \mathsf{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 3/4 \end{bmatrix}
\end{aligned}$$

Since u_1 and u_2 are orthogonal we have that

$$\mathsf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}; \mathsf{e}_2 = \begin{bmatrix} 0 \\ 4/5 \\ 3/5 \end{bmatrix}$$

For $\lambda = 50$, equation $\mathsf{Ax} - \lambda\mathsf{x} = 0$ results in $x_1 = 0, x_2 = -1$, and $x_3 = 3/4$ so that

$$\begin{aligned}
\mathsf{u}_3 &= \begin{bmatrix} 0 \\ 3/4 \\ -1 \end{bmatrix} \\
\mathsf{e}_3 &= \begin{bmatrix} 0 \\ 3/5 \\ -4/5 \end{bmatrix}
\end{aligned}$$

The corresponding modal matrix is given by

$$\mathsf{Q} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4/5 & 3/5 \\ 0 & 3/5 & -4/5 \end{bmatrix}$$

and the new coordinates are given by

$$\mathsf{x}' = \mathsf{Q}^{\mathsf{T}}\mathsf{x} = \begin{bmatrix} x_1 \\ 4/5x_2 + 3/5x_3 \\ 3/5x_2 - 4/5x_3 \end{bmatrix}$$

so that the canonical form of A becomes

$$\begin{aligned} A &= \mathsf{x}'^{\mathsf{T}}\mathsf{A}'\mathsf{x}' = \mathsf{x}'^{\mathsf{T}}(\mathsf{Q}^{\mathsf{T}}\mathsf{A}\mathsf{Q})\mathsf{x}' \\ &= 25x_1'^2 + 25x_2'^2 + 50x_3'^2 \end{aligned}$$

In particular, the surface $25x_1'^2 + 25x_2'^2 + 50x_3'^2 = 25$ or $x_1'^2 + x_2'^2 + 2x_3'^2 = 1$ represents an oblate spheroid.

3 Equivalent Matrices and Transformations

Definition 15 Two matrices A and B which can be obtained from each other by a finite number of successive applications of elementary operations to rows and/or columns are said to be equivalent (but not necessary equal) matrices.

Since any sequence of operations on the rows of A can be achieved by premultiplying A by some nonsingular matrix P, corresponding operations on columns can be always effected by postmultiplying A by a nonsingular matrix Q.

Therefore, the elementary operations on the rows of A are achieved as

$$\begin{aligned} \mathsf{PA} &= \mathsf{P}\begin{bmatrix} \mathsf{r}_1 \\ \vdots \\ \mathsf{r}_n \end{bmatrix} \\ & \begin{bmatrix} p_{11}\mathsf{r}_1 + \ldots + p_{1n}\mathsf{r}_n \\ \vdots \\ p_{n1}\mathsf{r}_1 + \ldots + p_{nn}\mathsf{r}_n \end{bmatrix} \end{aligned}$$

and similarly, elementary operations on the columns of A are accomplished by

$$\begin{aligned} \mathsf{AQ} &= [\mathsf{c}_1, \ldots, \mathsf{c}_n]\mathsf{Q} \\ &= \begin{bmatrix} q_{11}\mathsf{c}_1 + \ldots + q_{1n}\mathsf{c}_n \\ \vdots \\ q_{n1}\mathsf{c}_1 + \ldots + q_{nn}\mathsf{c}_n \end{bmatrix} \end{aligned}$$

Consider now and $m \times n$ matrix A. We perform elementary operations on an $m \times m$ identity matrix I. Say, (i) interchange two rows, (ii) multiply the elements of a row by a nonzero number and (iii) add to the elements of a row, k times the corresponding elements of another row.

Example 16 (i) For, say, row 1 and 2 we get

$$\mathsf{P} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\mathsf{PA} = \begin{bmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

(ii) Say, for the elementary operation to multiply k times the elements of the second row we have that

$$\mathsf{P} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\mathsf{PA} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ ka_{21} & ka_{22} & ka_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

(iii) For the elementary operation of multiplying row 1 by k and adding the resulting row to row 2 we get

$$\mathsf{P} = \begin{bmatrix} 1 & 0 & 0 \\ k & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\mathsf{PA} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ (ka_{11} + a_{21}) & (ka_{12} + a_{22}) & (ka_{13} + a_{23}) \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

(i') In order to switch columns one and two we have

$$\mathsf{Q} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\mathsf{AQ} = \begin{bmatrix} a_{12} & a_{11} & a_{13} \\ a_{22} & a_{21} & a_{23} \\ a_{32} & a_{31} & a_{33} \end{bmatrix}$$

etc.

Conclusion 17 Operation sequence on the rows of A is equivalent to performing PA, where P is obtained by performing the sequence of elementary operations on an identity matrix I. Operation sequence on columns of A is equivalent to AQ, where Q is obtained from I by performing on it corresponding elementary operations.

Theorem 18 Any nonsingular (square) matrix can be reduced to the unit matrix of the same order by use of only elementary row operations and also by use only of elementary column operations. Then deduce if $\mathsf{B} = \mathsf{PAQ}$, where $\det \mathsf{P} \neq 0$ and $\det \mathsf{Q} \neq 0$, then B can be obtained from A by use of elementary operations.

Proof. Let's consider the Gauss-Jordan procedure for a matrix A.

$$\begin{aligned}
\mathsf{A} &= \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \\
G - J &\Rightarrow \\
\mathsf{I} &= \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & & & \\ 0 & 0 & \dots & 1 \end{bmatrix}
\end{aligned}$$

All of these operations can be represented as elementary row operation sequence PA. Similarly for the column operations. The converse of the proceeding theorem is also true, i.e., through elementary operations A can be reduced to an identity matrix. Now the matrices A and B are equivalent iff nonsingular matrices P and Q exist such that $\mathsf{B} = \mathsf{PAQ}$. Since the elementary operations do not change the rank of a matrix we conclude that two equivalent matrices have the same rank. ■

Transformations of the form PAQ are classified according to restrictions imposed on P and Q as follows:

1. If $\mathsf{P} = \mathsf{Q}^\mathsf{T} = \mathsf{Q}^{-1} \Rightarrow \mathsf{Q}^{-1}\mathsf{AQ}$ is an orthogonal transformation.
2. If $\mathsf{P} = \mathsf{Q}^\mathsf{T} \Rightarrow \mathsf{Q}^\mathsf{T}\mathsf{AQ}$ is a congruence transformation, and
3. if $\mathsf{P} = \mathsf{Q}^{-1}$, $\mathsf{Q}^{-1}\mathsf{AQ}$ is a similarity transformation.