1 Linear independence

A subset $L \subseteq \mathbb{R}^n$ is called a *linear space*, if it is closed under $+$ and multiplication by a real number, that is,

- $\bar{x}, \bar{y} \in L \quad \Rightarrow \quad \bar{x} + \bar{y} \in L$;
- $\bar{x} \in L \quad \Rightarrow \quad \lambda\bar{x} \in L$, for all $\lambda \in \mathbb{R}$.

If $L_1 \subseteq L_2$ are linear spaces, then L_1 is called a *subspace* of L_2.

Examples. $\{\bar{0}\}$; $\mathbb{R}^n$; the set of vectors $\{\lambda\bar{a} : \lambda \in \mathbb{R}\}$, where $\bar{a}$ is any fixed vector.

Important example. The set of all solutions of a homogeneous system of linear equations $A\bar{x} = \bar{0}$ is a linear space.

Proof. We have to show that $L = \{\bar{x} : A\bar{x} = \bar{0}\}$ is closed under $+$ and $\lambda\,\cdot$. If $A\bar{x} = \bar{0}$ and $A\bar{y} = \bar{0}$, then $A(\bar{x} + \bar{y}) = A\bar{x} + A\bar{y} = \bar{0}$, using properties of matrix multiplication. Similarly, $A(\lambda\bar{x}) = \lambda A\bar{x} = \bar{0}$.

Definition. *Linear span* of vectors $\bar{x}_1, \ldots, \bar{x}_k$ is the minimal linear space L containing all of them. It can also be defined as the set of all *linear combinations* of the form $\lambda_1\bar{x}_1 + \cdots + \lambda_k\bar{x}_k$ for $\lambda_1, \ldots, \lambda_k \in \mathbb{R}^n$. (Check that it is indeed a linear space!) Denoted $L = \langle\bar{x}_1, \ldots, \bar{x}_k\rangle$.

Definition. $\bar{x}_1, \ldots, \bar{x}_k$ are *linearly dependent* if there is a nontrivial linear combination of them, which is equal $\bar{0}$:

$$\lambda_1\bar{x}_1 + \cdots + \lambda_k\bar{x}_k = \bar{0},$$

where not all $\lambda_i = 0$. Otherwise, $\bar{x}_1, \ldots, \bar{x}_k$ are called *linearly independent.*[1]

Examples.

1. Any set of vectors containing $\bar{0}$ is linearly dependent.
2. If $\bar{x} \neq \bar{0}$, then $\{\bar{x}\}$ is linearly independent.

[1] In the list $x_1, \ldots, x_k$ some vectors may occur several times in which case they are linearly dependent (why?). So, the notion of linear dependence really applies to *multisets* of vectors.

Important example. Consider n vectors $\bar{e}_i = (0 \ldots 010 \ldots 0)$, where 1 occurs in position i. These are linearly independent.

Proof: consider a linear combination

$$\lambda_1 \bar{e}_1 + \cdots + \lambda_n \bar{e}_n = \bar{0}.$$

Look at the i-th component of the vector on the left hand side. It is the sum of the i-th components of vectors $\lambda_j \bar{e}_j$. But i-th component of $\bar{e}_j$ equals 0, unless $j = i$, so only $\lambda_i \bar{e}_i$ remains, whose i-th component is λ_i. But on the right hand side we have $\bar{0}$, whose i-th component is 0, so $\lambda_i = 0$. This can be done for any i, so all coefficients have to be 0.

Definition. *Dimension* $\dim(L)$ of a vector space L is the maximal number k such that there are k linearly independent vectors in L.

Examples. $\dim(\{\bar{0}\}) = 0$; $\dim(\langle \bar{x} \rangle) = 1$, if $\bar{x} \neq \bar{0}$; $\dim(\mathbb{R}^n) = n$. (We already know n linearly independent vectors $\bar{e}_i$ in $\mathbb{R}^n$. We shall see below that there cannot be more than n linearly independent vectors in $\mathbb{R}^n$.)

Problem. Given k arbitrary vectors, how to check if they are linearly independent?

Solution: Write the given vectors as columns in a matrix $X = (\bar{x}_1 | \cdots | \bar{x}_k)$. Consider a vector $\bar{\lambda}$ of k (unknown) coefficients $\lambda_1, \ldots, \lambda_k$. We have in matrix notation

$$X\bar{\lambda} = \lambda_1 \bar{x}_1 + \cdots + \lambda_k \bar{x}_k,$$

hence the vectors are linearly independent if and only if the system of linear equations $X\bar{\lambda} = \bar{0}$ has only the trivial solution $\bar{\lambda} = \bar{0}$. But we can solve systems of linear equations by Gaussian elimination method. Bring the matrix X to a staircase form and examine, if all of its columns are basis columns (corresponding to basis variables of the associated system).

Example. Consider from this point of view the case of $n+1$ vectors in $\mathbb{R}^n$. Matrix X will have $n+1$ column and only n rows. So, there cannot be more than n "steps" in the staircase form of it. "Steps" correspond precisely to basis variables, therefore there has to be at least one parametric variable! Conclusion: $n+1$ vectors in $\mathbb{R}^n$ cannot be linearly independent, $\dim(\mathbb{R}^n) = n$.

2 Rank of a matrix

Definition. *Column rank* of a matrix A is defined as the maximum number k such that there are k linearly independent vectors among its columns. *Row rank* of A is defined similarly. Denoted $rk_{\rm col}(A)$ and $rk_{\rm row}(A)$, respectively.

Theorem 1 *For any matrix A, $rk_{col}(A) = rk_{row}(A)$.*

Proof. The proof consists of 4 steps:

1. Elementary transformations preserve column rank.
2. Elementary transformations preserve row rank.
3. One can bring any matrix by elementary transformations to a staircase form.
4. For a staircase matrix both ranks are equal.

We already know 3.

Proof of 1. We know that elementary transformations preserve the set of solutions of the system of linear equations with a given matrix. Consider a matrix A' corresponding to a maximal linearly independent set of columns of A.[2] The system $A'\bar{\lambda} = \bar{0}$ has only a trivial solution (by linear independence). Hence, the transformed system has the same property, and the corresponding columns of the transformed matrix A have to be linearly independent, too. Therefore, column rank does not decrease. But it also cannot increase, because elementary transformations are invertible. (Otherwise, the rank would decrease under the inverse transformation.)

Proof of 2. Again, it is sufficient to show that the rank does not decrease. The transformation of type $\bar{a} \mapsto \lambda\bar{a}$, where $\lambda \neq 0$, obviously maps a linearly independent set of vectors (containing or not containing $\bar{a}$) to a linearly independent one. The claim is also obvious for the transposition of two rows, because the set of rows remains the same.

Consider now the transformation

$$\bar{a} \mapsto \bar{a} + \bar{b}$$

applied to a pair of rows of A.

We need a little lemma.

[2] That is, any linearly independent set of $rk_{\rm col}(A)$ columns.

Lemma 1 *Assume $\bar{x}_1, \ldots, \bar{x}_k$ are linearly independent vectors, but $y, \bar{x}_1, \ldots, \bar{x}_k$ are linearly dependent. Then $y \in \langle \bar{x}_1, \ldots, \bar{x}_k \rangle$, that is, $\bar{y}$ equals a linear combination of $\bar{x}_i$.*

Proof. Assume

$$\lambda_0 \bar{y} + \lambda_1 \bar{x}_1 + \cdots + \lambda_k \bar{x}_k = \bar{0}.$$

If $\lambda_0 = 0$, we also obtain $\lambda_i = 0$ for $i \geq 1$ from the linear independence of $\bar{x}_1, \ldots, \bar{x}_k$. Otherwise, we have

$$\bar{y} = -\frac{\lambda_1}{\lambda_0} \bar{x}_1 - \cdots - \frac{\lambda_k}{\lambda_0} \bar{x}_k,$$

q.e.d.

Consider now a maximal linearly independent set of rows $\bar{a}_1, \ldots, \bar{a}_k$ in A. If $\bar{a}$ does not occur in this set, the same rows will occur in the transformed matrix, so the rank does not decrease.

So, assume that $\bar{a}$ occurs in this set, say $\bar{a} = \bar{a}_1$. Consider the row $\bar{b}$. If $\bar{b}, \bar{a}_2, \ldots, \bar{a}_k$ are linearly independent, we are done, because $\bar{b}$ is also a row of the transformed matrix (we found k linearly independent rows).

If this is not the case, $\bar{b} \in L := \langle \bar{a}_2, \ldots, \bar{a}_k \rangle$ (use the lemma). Now either $\bar{a}+\bar{b}, \bar{a}_2, \ldots, \bar{a}_k$ are linearly independent, and we are done for the same reason as above, or they are linearly dependent and then $\bar{a} + \bar{b} \in L$. Together with $\bar{b} \in L$ this implies $\bar{a} = (\bar{a} + \bar{b}) - \bar{b} \in L$ contradicting the assumption that $\bar{a}, \bar{a}_2, \ldots, \bar{a}_k$ were linearly independent.

Proof of 4. In a staircase matrix (simplified by Jordan elimination) both ranks equal the number of "steps" in the stairs: all (say, k) nonzero raws form a maximal linearly independent set. Basis columns are just the columns of the form $\bar{e}_i$, for $i \leq k$. As we have seen, they are linearly independent.

We only need to see that there cannot be more than k linearly independent columns. We proceed as follows. A staircase matrix can be simplified even further by elementary transformations with its *columns*. These change neither column nor row ranks, because they are equivalent to the elementary row transformations with the transposed matrix. Using elementary column transformations one can bring the staircase matrix to a form where each row is either a zero vector, or is one of the vectors $\bar{e}_i$, for $i \leq k$. A set of more then k columns of such a matrix would always contain a zero vector thus being linearly dependent, q.e.d.

Thus, rank of matrix makes sense independently of whether we consider rows or columns and can be denoted $rk(A) = rk_{\text{row}}(A) = rk_{\text{col}}(A)$.

Transposition of a matrix maps columns to rows and vice versa, hence we also have

Corollary 2 *$rk(A) = rk(A^T)$.*

3 Bases and coordinates

Definition. A list of vectors $\bar{x}_1, \ldots, \bar{x}_k$ is called a *basis* in a vector space L if the following conditions hold:

1. $\langle \bar{x}_1, \ldots, \bar{x}_k \rangle = L$;
2. $\bar{x}_1, \ldots, \bar{x}_k$ are linearly independent.

Theorem 2 *If $\bar{x}_1, \ldots, \bar{x}_k$ is a basis in L, then for every vector $\bar{y} \in L$ there are uniquely defined $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ such that*

$$\bar{y} = \lambda_1 \bar{x}_1 + \cdots + \lambda_k \bar{x}_k.$$

These coefficients are called coordinates of $\bar{y}$ with respect to the given basis.

Proof. Existence follows from Condition 1. If there were another linear combination

$$\bar{y} = \lambda'_1 \bar{x}_1 + \cdots + \lambda'_k \bar{x}_k,$$

we would have (subtracting the second from the first)

$$\bar{0} = (\lambda_1 - \lambda'_1)\bar{x}_1 + \cdots + (\lambda_k - \lambda'_k)\bar{x}_k.$$

Condition 2 then implies $\lambda_i - \lambda'_i = 0$, for all i, q.e.d.

Bases are usually given by matrices whose colums are the basis vectors, which we write as $(\bar{x}_1 | \cdots | \bar{x}_k)$.

Important example. The standard basis in $\mathbb{R}^n$ consists of vectors $\bar{e}_i$, $i = 1, \ldots, n$ introduced above. The corresponding matrix is the unit matrix E . Coordinates of any vector w.r.t. the standard basis are just its components.

Example. The vectors $\bar{x}_1 = (1,2)$ and $\bar{x}_2 = (1,3)$ form a basis in $\mathbb{R}^2$. Indeed, they are linearly independent. Yet, together with any other vector $\bar{y}$ they form a linearly dependent system, because $\dim(\mathbb{R}^2) = 2$. Hence by Lemma 1 $\bar{y} \in \langle \bar{x}_1, \bar{x}_2 \rangle$.

In a similar manner it is easy to see that, if $\dim(L) = n$, then any set of n linearly independent vectors in a vector space L is a basis.

Problem. Given a basis and a vector $\bar{x}$ in L find its coordinates.

Solution: assume a basis in L is given by its matrix X. The coordinate vector $\bar{\lambda}$ satisfies the following equation: $X\bar{\lambda} = \bar{x}$. To find $\bar{\lambda}$ solve the corresponding system of linear equations. (Its unique solvability precisely means that the coordinates are uniquely defined.)

Theorem 3 *If* $\dim(L) = k$*, then any basis in* L *has exactly* k *elements.*

Proof. It cannot have more than k elements, because those basis vectors have to be linearly independent. Assume the basis has $m < k$ elements, given by an $n \times m$ matrix X, and consider the matrix whose columns are the coordinates of a maximal linearly independent system of k vectors $\bar{y}_1, \ldots, \bar{y}_k$ (which exists because $dim(L) = k$) in the given basis. This is an $m \times k$ matrix, say A. By the solution of the previous Problem, we have $X\bar{a}_i = \bar{y}_i$, where $\bar{a}_i$ is the i-th column of A. In matrix notation this can be written as $XA = Y$, where $Y = (\bar{y}_1 \mid \ldots \mid \bar{y}_k)$.

The column rank of A is equal to its row rank, which is $\leq m$. But then the colums of A have to be linearly dependent, that is, a nontrivial solution $A\bar{\lambda} = \bar{0}$ exists. Then the same λ satisfies

$$(XA)\bar{\lambda} = X(A\bar{\lambda}) = X\bar{0} = \bar{0}.$$

Hence we have $Y\bar{\lambda} = \bar{0}$, which gives a nontrivial linear combination of the linearly independent set of k vectors, a contradiction, q.e.d.

Consider a special case of $\mathbb{R}^n$. Any linearly independent set of n vectors forms a basis in $\mathbb{R}^n$. The corresponding matrix X is a $n \times n$ matrix of rank n. Such matrices are called *regular*. If $rk(X) < n$, it is called *singular*.

Changing a basis. If you go from one basis, say $\bar{e}_1, \ldots, \bar{e}_n$, to another basis $\bar{e}'_1, \ldots, \bar{e}'_n$, you can uniquely define the corresponding $n \times n$ *transfer* *matrix* S, whose columns are the coordinates of the new basis vectors w.r.t. the old one. It satisfies the following matrix equation:

$$(\bar{e}'_1|\cdots|\bar{e}'_n) = (\bar{e}_1|\cdots|\bar{e}_n)S.$$

Also notice that matrices $(\bar{e}_1|\cdots|\bar{e}_n)$ and $(\bar{e}'_1|\cdots|\bar{e}'_n)$ can be thought of as the transfer matrices from the standard basis to the given ones.

If you go now from $\bar{e}'_1,\ldots,\bar{e}'_n$ to a third basis $\bar{e}''_1,\ldots,\bar{e}''_n$ with a transfer matrix S', you obtain

$$(\bar{e}''_1|\cdots|\bar{e}''_n) = (\bar{e}'_1|\cdots|\bar{e}'_n)S',$$

and hence

$$(\bar{e}''_1|\cdots|\bar{e}''_n) = (\bar{e}_1|\cdots|\bar{e}_n)SS'.$$

Thus, SS' is the transfer matrix from the first to the third basis. We have shown that transfer matrices multiply, if one changes basis several times.

Important problem. What is the relationship between the coordinates of a vector in the first and the second basis?

Solution: Coordinates $\bar{\lambda}'$ of a vector $\bar{x}$ in basis $\bar{e}'_1,\ldots,\bar{e}'_n$ satisfy $E_2\bar{\lambda}' = \bar{x}$, where $E_2 = (\bar{e}'_1|\cdots|\bar{e}'_n)$ is the matrix of the second basis. We have a similar equation for the coordinates $\bar{\lambda}$ of $\bar{x}$ in the first basis. But E_2 is the product of the matrix of the first basis and the transfer matrix: $E_2 = E_1 S$. Therefore,

$$E_1 S\bar{\lambda}' = \bar{x} = E_1\bar{\lambda},$$

that is, $\bar{\lambda} = S\bar{\lambda}'$ (coordinates are uniquely defined).

4 Inverse matrix

A matrix B is called *inverse* to A, if $AB = BA = E$. This is only possible, if A and B are square $n\times n$ matrices.

Not every square matrix has an inverse (think about zero matrix), but if the inverse exists, then it is uniquely defined.

Proof. If $AB_1 = B_1A = E$ and $AB_2 = B_2A = E$, then

$$B_1 = B_1E = B_1(AB_2) = (B_1A)B_2 = EB_2 = B_2.$$

The inverse to A, if exists, is denoted A^{-1}. If the inverse exists, A is called *invertible*.

Theorem 4 *A is invertible if and only if A is regular.*

Proof. Assume A is invertible. To show that A is regular we invoke an independently useful lemma.

Lemma 3 $rk(BA) \leq rk(A)$.

Proof. Notice that any linear dependence among the columns of A with coefficients $\bar{\lambda}$ is expressed in matrix form by the equation $A\bar{\lambda} = \bar{0}$. For such $\bar{\lambda}$ we also have

$$(BA)\bar{\lambda} = B(A\bar{\lambda}) = B\bar{0} = \bar{0}.$$

This means that the same linear dependence holds for the columns of the matrix BA. Hence, its rank cannot exceed that of A, q.e.d.

From this lemma, assuming B is the inverse of A, we can conclude

$$n = rk(E) = rk(BA) \leq rk(A) \leq n,$$

hence $rk(A) = n$.

Assume A is regular, then its columns form a basis $\bar{a}_1, \dots, \bar{a}_n$ in $\mathbb{R}^n$. Let $\bar{e}_1, \dots, \bar{e}_n$ be the standard basis. Consider the transfer matrix X from $\bar{a}_1, \dots, \bar{a}_n$ to $\bar{e}_1, \dots, \bar{e}_n$. Since transfer matrices multiply, we have $AX = E$ (go from the standard basis to the basis of column vectors of A and back). Similarly, we have $XA = E$ (go from the basis of colums of A to the standard one and back), q.e.d.

Problem. How to calculate the inverse of a given matrix A?

Solution. Column vectors $\bar{x}_i$ of A^{-1} satisfy the following systems of linear equations:

$$\begin{aligned} A\bar{x}_1 &= \bar{e}_1 \\ A\bar{x}_2 &= \bar{e}_2 \\ &\dots \\ A\bar{x}_n &= \bar{e}_n \end{aligned}$$

Here $\bar{e}_1, \dots, \bar{e}_n$ are the standard basis vectors. Calculating solutions can be done by bringing A to a unit matrix form by elementary transformations (familiar Gauss–Jordan elimination), while doing the same transformations to a unit matrix. Transformations of the unit matrix simultaneously represent those of all the right hand side vectors. The transformed unit matrix will be the inverse of A.

5 Bilinear and quadratic forms

Bilinear form is a function $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ satisfying the following conditions, for any vectors $\bar{x}, \bar{y}, \bar{x}', \bar{y}'$ and numbers λ:

1. $f(\lambda\bar{x}, \bar{y}) = \lambda f(\bar{x}, \bar{y})$;
2. $f(\bar{x} + \bar{x}', \bar{y}) = f(\bar{x}, \bar{y}) + f(\bar{x}', \bar{y})$;
3. $f(\bar{x}, \lambda\bar{y}) = \lambda f(\bar{x}, \bar{y})$;
4. $f(\bar{x}, \bar{y} + \bar{y}') = f(\bar{x}, \bar{y}) + f(\bar{x}, \bar{y}')$.

Conditions 1 and 2 are called 'linearity in the first argument'.

f is called *symmetric*, if $f(\bar{x}, \bar{y}) = f(\bar{y}, \bar{x})$. f is called *skew-symmetric*, if $f(\bar{x}, \bar{y}) = -f(\bar{y}, \bar{x})$.

Given a basis $\bar{e}_1, \ldots, \bar{e}_n$ (not necessarily standard), we can associate with a bilinear form its *matrix* $A = (a_{ij})$ such that $a_{ij} = f(\bar{e}_i, \bar{e}_j)$ for $1 \leq i, j \leq n$. In other words, elements of the matrix are the values of f at all possible pairs of basis vectors.

Let us see that this matrix uniquely determines the form. Indeed, if $\bar{x} = \sum_{i=1}^{n} \lambda_i \bar{e}_i$ and $\bar{y} = \sum_{j=1}^{n} \mu_j \bar{e}_j$, then

$$\begin{aligned} f(\bar{x}, \bar{y}) &= f\left(\sum_i \lambda_i \bar{e}_i, \sum_j \mu_j \bar{e}_j\right) \\ &= \sum_i f(\lambda_i \bar{e}_i, \sum_j \mu_j \bar{e}_j) \\ &= \sum_i \sum_j f(\lambda_i \bar{e}_i, \mu_j \bar{e}_j) \\ &= \sum_{ij} \lambda_i \mu_j f(\bar{e}_i, \bar{e}_j) \\ &= \sum_{ij} a_{ij} \lambda_i \mu_j. \end{aligned}$$

This can be expressed as follows:

$$f(\bar{x}, \bar{y}) = \bar{\lambda}^T A \, \bar{\mu},$$

where $\bar{\lambda}$, $\bar{\mu}$ are the coordinate vectors of $\bar{x}$, $\bar{y}$ in the given basis, and A is the matrix of f. Recalling that coordinates of a vector w.r.t. the standard basis are just its components, we can simply write

$$f(\bar{x}, \bar{y}) = \bar{x}^T A\, \bar{y},$$

if A is the matrix of the form in the standard basis.

Also notice that for any $n \times n$ matrix A the previous equation defines a bilinear function (check it!). So, there is a canonical correspondence between matrices and bilinear forms, provided a basis is fixed.

How does the matrix of a form change if one changes basis? Let S be the transfer matrix, $\bar{x}', \bar{y}'$ be the coordinates of $\bar{x}, \bar{y}$ in the first basis, and $\bar{x}'', \bar{y}''$ in the second basis, respectively. We have $\bar{x}' = S\bar{x}''$ and $\bar{y}' = S\bar{y}''$. Hence,

$$\bar{x}'^T A\bar{y}' = (S\bar{x}'')^T A(S\bar{y}'') = \bar{x}''^T (S^T AS)\bar{y}''.$$

Since the matrix of a bilinear form in a given basis is uniquely defined, $S^T AS$ must be the matrix of f in the second basis.

Example. Let a bilinear form be given in the standard basis by matrix $A = \begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix}$. Write its analytic expression in coordinates x_1, x_2, y_1, y_2.

Answer: Calculating $(x_1, x_2)A \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ yields

$$f(\bar{x}, \bar{y}) = x_1y_1 + 3x_2y_1 + 2x_1y_2 + 5x_2y_2.$$

Definition. *Quadratic form* is a function of the form

$$q(\bar{x}) = f(\bar{x}, \bar{x}),$$

where f is a symmetric bilinear form.

A quadratic form uniquely determines the associated symmetric bilinear form, which can be recovered by

$$f(\bar{x}, \bar{y}) = \frac{1}{2}(q(\bar{x} + \bar{y}) - q(\bar{x}) - q(\bar{y})).$$

Matrix of a quadratic form is the matrix of the corresponding bilinear form, thus in a given basis we have

$$q(\bar{x}) = \bar{\lambda}^T A\bar{\lambda},$$

where $\bar{\lambda}$ are the coordinates of $\bar{x}$.

A basis is *canonical* for a quadratic form, if the form's matrix is diagonal and the string of diagonal elements a_{ii} has the form $-1 \cdots -1 1 \cdots 1 0 \cdots 0$ (where some of -1's, 1's and 0's may be missing).

Theorem 5 *For every quadratic form there is a canonical basis.*

The canonical basis need not be uniquely defined, the point is that at least one exists! However, the matrix of the form in any canonical basis will be the same. We shall show it later.

Proof (idea). You can do it with elementary transformations, if you apply each elementary transformation symmetrically to the rows and the columns. This cooresponds to applying to A the operations $S^T AS$, for some (elementary) regular matrices S, q.e.d.

Typical excercise. Bring the quadratic form

$$\begin{pmatrix} 1 & 2 & -1 \\ 2 & 4 & 1 \\ -1 & 1 & 1 \end{pmatrix}$$

to a canonical basis and find the corresponding transfer matrix.

Solution. To find the transfer matrix, the elementary transformations applied to the rows of the matrix should also be applied to the unit matrix. (The symmetrical transformations with the columns are ignored.) The resulting matrix will be the *transposition* of the needed transfer matrix.

$$\begin{pmatrix}1&2&-1\\2&4&1\\-1&1&1\end{pmatrix}\quad\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}\qquad \bar{a}_2\mapsto\bar{a}_2-2\bar{a}_1 \tag{1}$$

$$\begin{pmatrix}1&2&-1\\0&0&3\\-1&1&1\end{pmatrix}\quad\begin{pmatrix}1&0&0\\-2&1&0\\0&0&1\end{pmatrix}\qquad \bar{a}^2\mapsto\bar{a}^2-2\bar{a}^1 \tag{2}$$

$$\begin{pmatrix}1&0&-1\\0&0&3\\-1&3&1\end{pmatrix}\quad\begin{pmatrix}1&0&0\\-2&1&0\\0&0&1\end{pmatrix}\qquad \bar{a}_3\mapsto\bar{a}_3+\bar{a}_1 \tag{3}$$

$$\begin{pmatrix}1&0&-1\\0&0&3\\0&3&0\end{pmatrix}\quad\begin{pmatrix}1&0&0\\-2&1&0\\1&0&1\end{pmatrix}\qquad \bar{a}^3\mapsto\bar{a}^3+\bar{a}^1 \tag{4}$$

$$\begin{pmatrix}1&0&0\\0&0&3\\0&3&0\end{pmatrix}\quad\begin{pmatrix}1&0&0\\-2&1&0\\1&0&1\end{pmatrix}\qquad \bar{a}_2\mapsto\bar{a}_2+\bar{a}_3 \tag{5}$$

$$\begin{pmatrix}1&0&0\\0&3&3\\0&3&0\end{pmatrix}\quad\begin{pmatrix}1&0&0\\-1&1&1\\1&0&1\end{pmatrix}\qquad \bar{a}^2\mapsto\bar{a}^2+\bar{a}^3 \tag{6}$$

$$\begin{pmatrix}1&0&0\\0&6&3\\0&3&0\end{pmatrix}\quad\begin{pmatrix}1&0&0\\-1&1&1\\1&0&1\end{pmatrix}\qquad \bar{a}_3\mapsto\bar{a}_3-(1/2)\bar{a}_2 \tag{7}$$

$$\begin{pmatrix}1&0&0\\0&6&3\\0&0&-\frac{3}{2}\end{pmatrix}\quad\begin{pmatrix}1&0&0\\-1&1&1\\3/2&-1/2&1/2\end{pmatrix}\qquad \bar{a}^3\mapsto\bar{a}^3-(1/2)\bar{a}^2 \tag{8}$$

$$\begin{pmatrix}1&0&0\\0&6&0\\0&0&-\frac{3}{2}\end{pmatrix}\quad\begin{pmatrix}1&0&0\\-1&1&1\\3/2&-1/2&1/2\end{pmatrix}\qquad \bar{a}_2\mapsto(1/\sqrt{6})\bar{a}_2 \tag{9}$$

$$\begin{pmatrix}1&0&0\\0&\sqrt{6}&0\\0&0&-\frac{3}{2}\end{pmatrix}\quad\begin{pmatrix}1&0&0\\-\frac{1}{\sqrt{6}}&\frac{1}{\sqrt{6}}&\frac{1}{\sqrt{6}}\\3/2&-1/2&1/2\end{pmatrix}\qquad \bar{a}^2\mapsto(1/\sqrt{6})\bar{a}^2 \tag{10}$$

$$\begin{pmatrix}1&0&0\\0&1&0\\0&0&-\frac{3}{2}\end{pmatrix}\quad\begin{pmatrix}1&0&0\\-\frac{1}{\sqrt{6}}&\frac{1}{\sqrt{6}}&\frac{1}{\sqrt{6}}\\3/2&-1/2&1/2\end{pmatrix}\qquad \begin{cases}\bar{a}_3\mapsto\frac{\sqrt{2}}{\sqrt{3}}\bar{a}_3\\ \bar{a}^3\mapsto\frac{\sqrt{2}}{\sqrt{3}}\bar{a}^3\end{cases} \tag{11}$$

$$\begin{pmatrix}1&0&0\\0&1&0\\0&0&-1\end{pmatrix} \quad \begin{pmatrix}1&0&0\\-\frac{1}{\sqrt{6}}&\frac{1}{\sqrt{6}}&\frac{1}{\sqrt{6}}\\\frac{\sqrt{3}}{\sqrt{2}}&-\frac{1}{\sqrt{6}}&\frac{1}{\sqrt{6}}\end{pmatrix} \quad \begin{cases}\bar{a}_3 \mapsto \bar{a}_1\\ \bar{a}_1 \mapsto \bar{a}_3\\ \bar{a}^3 \mapsto \bar{a}^1\\ \bar{a}^1 \mapsto \bar{a}^3\end{cases} \tag{12}$$

$$\begin{pmatrix}-1&0&0\\0&1&0\\0&0&1\end{pmatrix} \quad \begin{pmatrix}\frac{\sqrt{3}}{\sqrt{2}}&-\frac{1}{\sqrt{6}}&\frac{1}{\sqrt{6}}\\-\frac{1}{\sqrt{6}}&\frac{1}{\sqrt{6}}&\frac{1}{\sqrt{6}}\\1&0&0\end{pmatrix} \quad \text{That's all} \tag{13}$$

So, the resulting transfer matrix (to the canonical basis) is

$$S = \begin{pmatrix}\frac{\sqrt{3}}{\sqrt{2}}&-\frac{1}{\sqrt{6}}&1\\-\frac{1}{\sqrt{6}}&\frac{1}{\sqrt{6}}&0\\\frac{1}{\sqrt{6}}&\frac{1}{\sqrt{6}}&0\end{pmatrix}$$

Example. The following are all possible matrices of 2-dimensional quadratic forms in a canonical basis and the corrsponding analytic expressions (in coordinates $\bar{x} = (x_1, x_2)$):

(i) $\begin{pmatrix}1&0\\0&1\end{pmatrix}$, $q(\bar{x}) = x_1^2 + x_2^2$;

(ii) $\begin{pmatrix}-1&0\\0&-1\end{pmatrix}$, $q(\bar{x}) = -x_1^2 - x_2^2$;

(iii) $\begin{pmatrix}-1&0\\0&1\end{pmatrix}$, $q(\bar{x}) = -x_1^2 + x_2^2$;

(iv) $\begin{pmatrix}1&0\\0&0\end{pmatrix}$, $q(\bar{x}) = x_1^2$;

(v) $\begin{pmatrix}-1&0\\0&0\end{pmatrix}$, $q(\bar{x}) = -x_1^2$;

(vi) $\begin{pmatrix}0&0\\0&0\end{pmatrix}$, $q(\bar{x}) = 0$.

Matrices A of quadratic forms in a canonical basis are characterized by two numbers: their *rank* $rk(A)$ and *index* $ind(A) =$ the number of -1's in the matrix.

Theorem 6 *$rk(A)$ and $ind(A)$ are invariants, that is, do not depend on the choice of a (canonical) basis.*

As a corollary we obtain that the matrix of a form in any canonical basis should be the same: the number of -1's on the diagonal is determined by $ind(A)$ and the number of 0's by $rk(A)$, respectively.

Proof. For rank this is easy to see, because multiplying a matrix by a regular matrix does not change its rank: $rk(SA) \leq rk(A)$ by Lemma 3, and $rk(A) = rk(S^{-1}SA) \leq rk(SA)$ by the same lemma, hence $rk(SA) = rk(A)$. But we also have

$$rk(AS) = rk((AS)^T) = rk(S^T A^T) = rk(A^T) = rk(A).$$

It follows that for any transfer matrix S, $rk(S^T AS) = rk(A)$.

Index of a quadratic form can be characterized as the maximal dimension of a linear subspace $L \subseteq \mathbb{R}^n$ on which the form is *negative definite*:

$$\forall \bar{x} \in L \ (\bar{x} \neq \bar{0} \Rightarrow q(\bar{x}) < 0).$$

(The argument is not difficult, but we skip it.) Hence index is also invariant and we can write $ind(q)$ and $rk(q)$ instead of $rk(A)$ and $ind(A)$.

If $ind(q) = 0$ and $rk(q) = n$, we say that the form is *positive definite.*