Lecture 3: Calculus and Linear Algebra

Piotr Zwiernik

Mathematics Brush-up

Barcelona School of Economics

Chapter 7: Systems of Linear equations

Many problems in economics can be modeled as system of linear equations.

In previous chapters we have already encountered systems of linear equations.

In this chapter, we consider some basic properties of such systems and discuss general solution procedures.

Read Chapter 8 of Werner-Sotskov and Chapter 7 of Simon-Blume

Exercises: 8.5, 8.6, 8.8 (Werner-Sotskov)

Systems of linear equations

Let A be a $m \times n$ matrix, $\mathbf{x} = (x_1, \ldots, x_n)$ a vector (unknown) in $\mathbb{R}^n$ and $\mathbf{b}$ a vector (given) in $\mathbb{R}^m$. Then the equation

$$A\mathbf{x} = \mathbf{b}$$

is a system of m linear equations and n variables.

In vector representation, the system writes as

$$\sum_{j=1}^{n} x_j \mathbf{a_j} = \mathbf{b},$$

where $\mathbf{a_1}, \ldots, \mathbf{a_n}$ are the column vectors of A.

If $\mathbf{b} = \mathbf{0}$ the system is called homogeneous, otherwise is called inhomogeneous. An inhomogeneous system can have no solution (inconsistent system) or at least one solution (consistent system). An homogeneous system has always the solution $\mathbf{x} = \mathbf{0}$.

Existence and uniqueness of a solution

Theorem: Let A be a $m \times n$ matrix. The maximum number of linearly independent column vectors of A coincides with the maximum number of linearly independent row vectors.

This number is called the rank of A, and denoted $r(A)$.

Theorem: The rank of a matrix A equals to the largest order of a minor of A that is different from zero. Recall that a minor is the determinant of a square submatrix of A.

Example:

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 4 & 6 & 2 \\ 3 & 2 & 4 \end{pmatrix}$$

Since $|A| = 0$, $r(A) \leq 2$. Since $|A_{33}| = -2$, $r(A) = 2$.

Observe that $2\mathbf{a_1} - \mathbf{a_2} = \mathbf{a_3}$ but none of the columns is a multiple of the other, so the families $\{\mathbf{a_1}, \mathbf{a_2}\}$, $\{\mathbf{a_1}, \mathbf{a_3}\}$ and $\{\mathbf{a_2}, \mathbf{a_3}\}$ are l.i. and span the same subspace.

Computing the rank with Gaussian elimination

The rank of a matrix does not change either if we apply elementary columns or row operations to the matrix.

Example:

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 4 & 6 & 2 \\ 3 & 2 & 4 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 0 \\ 0 & -2 & 2 \\ 0 & -4 & 4 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 0 \\ 0 & -2 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

Since the third row has no pivot the columns are l.d. and since the matrix has 2 pivots the rank is 2. The pivot is the first non-zero value of a non-zero row after reducing the matrix with Gaussian elimination. The number of pivots equals the rank.

To find a relation between columns we need to solve the homogeneous system

$$\begin{aligned} \lambda_1 + 2\lambda_2 &= 0 \\ -\lambda_2 + \lambda_3 &= 0 \end{aligned}$$

The solution is $(-2\lambda_3, \lambda_3, \lambda_3)$. If for example $\lambda_3 = 1$, we obtain the relation

$$-2\mathbf{a_1} + \mathbf{a_2} + \mathbf{a_3} = 0.$$

Existence and uniqueness of a solution

We define the augmented matrix of A the $m \times (n+1)$ matrix whose columns are the columns of A plus the vector $\mathbf{b}$. We denote it by $A_{\mathbf{b}}$.

Remark: Either $r(A_{\mathbf{b}}) = r(A)$, either $r(A_{\mathbf{b}}) = r(A) + 1$.

Theorem: A linear system $A\mathbf{x} = \mathbf{b}$ is consistent (at least one solution) if and only if $r(A_{\mathbf{b}}) = r(A)$. (which is equivalent to $\mathbf{b} \in \mathrm{Span}\{\mathbf{a_1}, \ldots, \mathbf{a_n}\}$)

Theorem: Consider a consistent linear system $A\mathbf{x} = \mathbf{b}$, where A is $m \times n$.

1. If $r(A_{\mathbf{b}}) = r(A) = n$, then the solution is unique.

2. If $r(A_{\mathbf{b}}) = r(A) < n$, then there exist infinite many solutions and we need to choose $n - r(A)$ variables free.

Elementary transformation; solution procedures

Theorem: The set of solution does not change if we apply one of the elementary row operations to the matrix $A_{\mathbf{b}}$.

Theorem (Gauss elimination): Applying elementary row operations to the augmented matrix any system can be transformed into a triangular system.

Example:

$$\begin{aligned} x_1 + x_2 + x_3 &= 3 \\ x_1 - x_2 + 2x_3 &= 2 \\ 4x_1 + 6x_2 - x_3 &= 9 \end{aligned} \qquad \Longleftrightarrow \qquad \begin{aligned} x_1 + x_2 + x_3 &= 3 \\ -2x_2 + x_3 &= -1 \\ -4x_3 &= -4 \end{aligned}$$

We have done the row operations: $r_2 - r_1$, $r_3 - 4r_1$ and $r_3 + r_2$.

We have $r(A_{\mathbf{b}}) = r(A) = 3$ (3 pivots) so the solution is unique.

Both systems have the unique solution $(1, 1, 1)$.

Example

Consider the homogeneous system of equations

$$\begin{aligned} a+b+c+d+e+f &= 0 \\ 2a+2b+2c+2d-e-f &= 0 \\ 3a+3b-c-d-e-f &= 0 \end{aligned}$$

By Gaussian elimination, we obtain

$$A \sim \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

This matrix has 3 pivots so the rank is 3, so the set of solutions is a subspace of dimension $6 - r(A) = 3$.

This corresponds to the homogeneous system $a+b=0$, $c+d=0$, $e+f=0$. We choose 3 variables free and we obtain the general solution

$$(a, -a, c, -c, e, -e).$$

A basis of the subspace of solutions is $\{(1,-1,0,0,0,0),(0,0,1,-1,0,0),(0,0,0,0,1,-1)\}$.

General solution

Theorem: Every solution to the system $A\mathbf{x} = \mathbf{b}$ can be written as

$$\mathbf{x} = \mathbf{x}_h + \mathbf{x}_p,$$

where $\mathbf{x}_h$ is the general solution to the homogeneous system $A\mathbf{x} = \mathbf{0}$, and $\mathbf{x}_p$ is a particular solution of $A\mathbf{x} = \mathbf{b}$.

Corollary: If $A\mathbf{x} = \mathbf{b}$ is consistent, then the number of solutions is the same as the number of solutions to $A\mathbf{x} = \mathbf{0}$.

Example:

$$\begin{aligned} x_1 + x_2 + x_3 &= 1 \\ x_1 + x_2 + x_3 &= 1 \\ x_1 + x_2 + x_3 &= 1 \end{aligned} \qquad \Longleftrightarrow \qquad x_1 + x_2 + x_3 = 1$$

The matrix has 1 pivot so $r(A) = r(A_{\mathbf{b}}) = 1$, so we choose 2 variables free. General solution:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \lambda_1 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \lambda_1, \lambda_2 \in \mathbb{R}$$

Examples

1. Consider the matrix
$$A = \begin{pmatrix} 1 & 2 & 0 \\ 4 & 6 & 2 \\ 3 & 2 & 4 \end{pmatrix}$$
We have already seen that $r(A) = 2$, thus the solution to the system $A\mathbf{x} = \mathbf{0}$ has dimension $3 - r(A) = 1$ and is the line $\{\lambda(-2, 1, 1), \lambda \in \mathbb{R}\}$.
2. Consider the system of equations
$$\begin{aligned} x_1 + x_2 + x_3 &= -1 \\ x_1 + 2x_2 + 4x_3 &= 2 \end{aligned}$$
By Gauss elimination we obtain the equivalent system
$$\begin{aligned} x_1 + x_2 + x_3 &= -1 \\ x_2 + 3x_3 &= 3 \end{aligned}$$
Therefore $r(A) = r(A_{\mathbf{b}}) = 2$ and we choose 1 variable free. The solution is $\{\lambda(2, -3, 1) + (-4, 3, 0), \lambda \in \mathbb{R}\}$.

Example

We want to compute the canonical basis of the subspace spanned by the vectors

$$\mathbf{v_1} = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} \qquad \mathbf{v_2} = \begin{pmatrix} 3 \\ 0 \\ -1 \end{pmatrix} \qquad \mathbf{v_3} = \begin{pmatrix} 0 \\ 3 \\ -2 \end{pmatrix}$$

We do Gaussian row elimination to the transpose matrix

$$\begin{pmatrix} 2 & -1 & 0 \\ 3 & 0 & -1 \\ 0 & 3 & -2 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & -1/3 \\ 0 & 1 & -2/3 \\ 0 & 0 & 0 \end{pmatrix}$$

If we de not consider the transpose matrix then we need to do column operations.

Doing row operations and solving the homogeneous system we obtain the relation $3\mathbf{v_1} - 2\mathbf{v_2} + \mathbf{v_3} = 0$

The canonical basis is $\{(1, 0, -1/3), (0, 1, -2/3)\}$.

Other bases are $\{\vec{v}_1, \vec{v}_2\}$, $\{\vec{v}_1, \vec{v}_3\}$ and $\{\vec{v}_2, \vec{v}_3\}$.

Chapter 8: Eigenvalue problems and quadratic forms

Eigenvalue problems and quadratic forms are useful for deciding whether a function has an extreme point or for solving certain types of differential equations.

Eigenvalues often arise in economic problems dealing with processes of proportionate growth or decline.

Quadratic forms play an important role in optimization problems of functions of several variables.

Read Chapter 10 of Werner-Sotskov and (optional) Chapter 7 of Simon-Blume

Exercises: 10.2, 10.5 A,B,C, 10.6 a), b) (Werner-Sotskov)

Eigenvalues and eigenvectors: Economic example

Let x_t^M and x_t^W be the number of men and women in some population at time t.

We assume they satisfy the relation

$$\begin{pmatrix} x_{t+1}^M \\ x_{t+1}^W \end{pmatrix} = \begin{pmatrix} 0.8 & 0.4 \\ 0.3 & 0.9 \end{pmatrix} \begin{pmatrix} x_t^M \\ x_t^W \end{pmatrix}$$

Moreover, we assume the ratio of men and woman is constant over time, that is,

$$\begin{pmatrix} x_{t+1}^M \\ x_{t+1}^W \end{pmatrix} = \lambda \begin{pmatrix} x_t^M \\ x_t^W \end{pmatrix}$$

Question: Do λ and $\mathbf{x_t}' = (x_t^M, x_t^W)$ satisfying the above equations exist ?

Eigenvalues and eigenvectors

Let A be an $n \times n$ matrix. A scalar $\lambda \in \mathbb{R}$ is called an eigenvalue of A if there exists $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \neq \mathbf{0}$ solution to the equation

$$A\mathbf{x} = \lambda \mathbf{x}.$$

Then $\mathbf{x}$ is called an eigenvector of A with eigenvalue λ.

Intuitively the linear transformation given by A transforms $\mathbf{x}$ into a co-linear vector.

By definition the vector $\mathbf{0}$ is never considered an eigenvector!

The example above corresponds to finding the eigenvalues and eigenvectors of the matrix

$$A = \begin{pmatrix} 0.8 & 0.4 \\ 0.3 & 0.9 \end{pmatrix}$$

Eigenvalues and eigenvectors

Remark:

$$A\mathbf{x} = \lambda\mathbf{x} \Longleftrightarrow (A - \lambda I)\mathbf{x} = \mathbf{0}$$

Theorem: λ is an eigenvalue of $A \Leftrightarrow \text{rank}(A - \lambda I) < n \Leftrightarrow \det(A - \lambda I) = 0$.

The equation

$$P(\lambda) = \det(A - \lambda I) = 0$$

is called characteristic equation of A, and it is a polynomial of degree n in λ. The eigenvalues of A are the roots of the characteristic equation. Therefore, A has at most n real eigenvalues (some can be complex!).

General procedure :

1. Find the eigenvalues of A by solving the characteristic equation.
2. For each eigenvalue solve the system $(A - \lambda I)\mathbf{x} = \mathbf{0}$ in order to find the eigenvectors.

Eigenvalues and eigenvectors: Economic example

Let

$$A = \begin{pmatrix} 0.8 & 0.4 \\ 0.3 & 0.9 \end{pmatrix}$$

The characteristic equation is $(0.8 - \lambda)(0.9 - \lambda) - 0.12 = 0$, that is $\lambda^2 - 1.7\lambda + 0.6 = 0$, whose solutions are

$$\lambda_1 = 1.2 \quad \text{and} \quad \lambda_2 = 0.5.$$

If $\lambda_1 = 1.2$, the eigenvectors are $\{x(1, 1), x \neq 0\}$.

If $\lambda_2 = 0.5$, the eigenvectors are $\{x(1, -3/4), x \neq 0\}$.
The second solution does no make sense for our population growth model.

Solution: yes, they exist, we must have $x_0^M = x_0^W$ and the population grows by 20% from t to $t + 1$.

Eigenvalues and eigenvectors

Theorem: Let A $n \times n$. Then

1. Eigenvectors associated with different eigenvalues are linearly independent.
2. If A is symmetric, then A has all eigenvalues real and eigenvectors associated with different eigenvalues are orthogonal.

Example: Let

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

Then $P(\lambda) = \lambda^2 - 2\lambda - 3$ so the eigenvalues are

$$\lambda_1 = 3 \quad \text{and} \quad \lambda_2 = -1.$$

If $\lambda_1 = 3$, the eigenvectors are $\{x(1,1), x \neq 0\}$.
If $\lambda_2 = -1$, the eigenvectors are $\{x(-1,1), x \neq 0\}$.

Repeated eigenvalues

Example: Let

$$A = \begin{pmatrix} 0 & -1 & 1 \\ -7 & 0 & 5 \\ -5 & -2 & 5 \end{pmatrix}$$

Then $P(\lambda) = (\lambda - 2)^2(\lambda - 1)$ so the eigenvalues are

$$\lambda_1 = 2 \quad \text{and} \quad \lambda_2 = 1.$$

We call λ_1 an eigenvalue of multiplicity 2. In this case, the set of eigenvectors can have dimension 1 or 2.

If $\lambda_1 = 2$, the eigenvectors are $\{x(2,1,3), x \neq 0\}$ (dimension 1).

If $\lambda_2 = 1$, the eigenvectors are $\{x(1,-1,1), x \neq 0\}$.

Application: Principal Component Analysis

Principal components are new variables that are constructed as linear combinations of the initial variables.

These combinations are done in such a way that these new variables are uncorrelated.

What measures the amount of information is the variance, and principal components can be geometrically seen as the directions of high-dimensional data which capture the maximum amount of variance and project it onto a smaller dimensional subspace while keeping most of the information.

It is used in order to reduce the dimensionality of a large data set.

Application: Principal Component Analysis

Consider d variables $(X_1, \ldots, X_d)$. A data set of these variables is set of n vectors $\mathbf{x}_1, \ldots, \mathbf{x}_n$ in $\mathbb{R}^d$.

For each variable X_j we define the sample mean by

$$\overline{x}_j = \frac{1}{n} \sum_{i=1}^{n} x_{ij}.$$

The sample covariance matrix of the data set is the $d \times d$ symmetric matrix $\Sigma = (s_{jk})$:

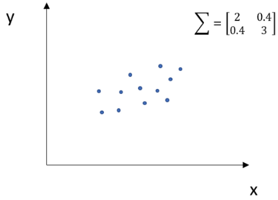
$$s_{jk} = \frac{1}{n-1} \sum_{i=1}^{n} (x_{ij} - \overline{x}_j)(x_{ik} - \overline{x}_k).$$

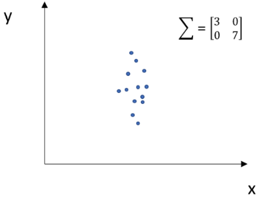
This matrix defines both the spread (variance) and the orientation (covariance) of our data.

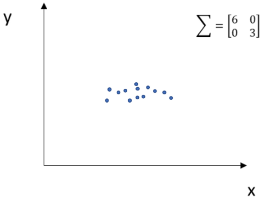
The sample covariance is an unbaised estimator of the covariance $E((X_j - E(X_j))(X_k - E(X_k))$

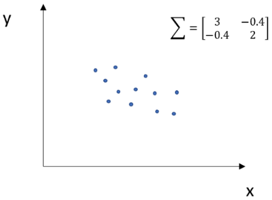
Application: Principal Component Analysis

Examples of covariance matrices with $d = 2$ and $n = 13$:



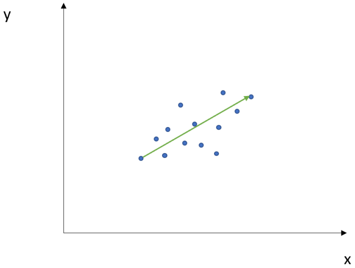






Application: Principal Component Analysis

To this matrix Σ we assign two elements: a vector that will point into the direction of the larger spread of data and a number equal to the spread (variance) of that direction. These two elements are, respectively, an Eigenvector and Eigenvalue of Σ. If we sort our eigenvectors in descending order with respect to their eigenvalues, we will have that the first eigenvector accounts for the largest spread among data, and so forth (under the condition that all these new directions, which describe a new space, are orthogonal among each other).



Numerical Example (PCA)

age	weight	height
20	50	4
25	60	5
30	65	5.5
40	75	6

Covariance matrix:

$$\begin{pmatrix} 72.9 & 87.5 & 6.875 \\ 87.5 & 108.3 & 8.75 \\ 6.875 & 8.75 & 0.73 \end{pmatrix}$$

Eigenvectors (PCA components)

$$\begin{pmatrix} 0.63 & 0.77 & 0.06 \\ 0.77 & -0.617 & -0.16 \\ 0.08 & -0.148 & 0.98 \end{pmatrix}$$

Eigenvalues (explained variances of the components)

$$(180, 1.38, 0.005)$$

Numerical Example (PCA)

If there are eigenvalues close to zero, they represent components that may be discarded.

We can compute the explained variance ratio of the first component which is 0.99, that is, the first component is enough to explain up to 99% variance in the data.

We can now project our data into a $4x1$ matrix instead of a $4x3$ matrix, thereby reducing the dimension of the variables, with a minor loss in information.

This is done multiplying the data matrix with the eigenvector

$$\begin{pmatrix} 20 & 50 & 4 \\ 25 & 60 & 5 \\ 30 & 65 & 5.5 \\ 40 & 75 & 6 \end{pmatrix} \begin{pmatrix} 0.63 \\ 0.77 \\ 0.08 \end{pmatrix} = \begin{pmatrix} 51.42 \\ 62.35 \\ 69.39 \\ 83.43 \end{pmatrix}$$

This can be seen as a linear regression where the regression coefficients are the components of the eigenvectors

Quadratic forms

A quadratic form is a function $Q : \mathbb{R}^n \to \mathbb{R}$ of the form

$$Q(\mathbf{x}) = \mathbf{x}^{\mathrm{T}} A \mathbf{x} = \sum_{i,j=1}^{n} a_{ij} x_i x_j,$$

where $A = (a_{ij})$ is a $n \times n$ symmetric matrix and $\mathbf{x}' = (x_1, \ldots, x_n)$ is a vector in $\mathbb{R}^n$.

Remark: Since the coefficient of $x_i x_j$ is $a_{ij} + a_{ji}$, the quadratic form doesn't change if we replace both coefficients by $\frac{a_{ij}+a_{ji}}{2}$, which corresponds to assume that A is symmetric.

Example: Both matrices

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 5/2 \\ 5/2 & 4 \end{pmatrix}$$

have the same quadratic form $Q(x, y) = x^2 + 4y^2 + 5xy$.

For each quadratic form there exists a unique symmetric matrix

Quadratic forms and their sign

A $n \times n$ symmetric matrix is said to be :

1. Positive definite if $\mathbf{x}'A\mathbf{x} > 0$, for all $\mathbf{x} \neq \mathbf{0}$.
2. Negative definite if $\mathbf{x}'A\mathbf{x} < 0$ for all $\mathbf{x} \neq \mathbf{0}$.
3. Positive semi-definite if $\mathbf{x}'A\mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$.
4. Negative semi-definite if $\mathbf{x}'A\mathbf{x} \leq 0$ for all $\mathbf{x} \in \mathbb{R}^n$.
5. Indefinite if both $\mathbf{x}'A\mathbf{x} > 0$ and $\mathbf{x}'A\mathbf{x} < 0$ are possible.

Example:

$$\mathbf{x}'A\mathbf{x} = (x, y)\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = (x - y)^2 \geq 0$$

Therefore, A is positive semi-definite.

Quadratic forms and their sign

Theorem: Let A a $n \times n$ symmetric matrix. Then:

1. A is positive definite $\iff$ all eigenvalues of A are > 0.
2. A is negative definite $\iff$ all eigenvalues of A are < 0.
3. A is positive semi-definite $\iff$ all eigenvalues of A are ≥ 0.
4. A is negative semi-definite $\iff$ all eigenvalues of A are ≤ 0.
5. A is indefinite $\iff$ A has two eigenvalues of different sign.

Example: The eigenvalues of $A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$ are $\lambda_1 = 0$ and $\lambda_2 = 2$ since $P(\lambda) = \lambda(\lambda - 2)$.

Quadratic forms and their sign

Let A be a $n \times n$ matrix. The principal minors are the determinants Δ_k obtained crossing out the same $n-k$ columns and rows of A. The leading principal minors are the determinants D_k obtained crossing out the last $n-k$ columns and rows of A.

Theorem: Let A a $n \times n$ symmetric matrix. Then:

1. A is positive definite $\iff D_k > 0$ for all $k = 1, ..., n$.
2. A is negative definite $\iff (-1)^k D_k > 0$ for all $k = 1, ..., n$.
3. If A is positive semi-definite $\iff \Delta_k \geq 0$ for all $k = 1, ..., n$.
4. If A is negative semi-definite $\iff (-1)^k \Delta_k \geq 0$ for all $k = 1, ..., n$.
5. If none of the above conditions hold then A is indefinite.

Example:

$$A = \begin{pmatrix} -3 & 2 & 0 \\ 2 & -3 & 0 \\ 0 & 0 & -5 \end{pmatrix}$$

Since $D_1 = -3$, $D_2 = 5$, and $D_3 = -25$, A is negative definite.

Examples

1. Consider the matrix
$$A = \begin{pmatrix} 1 & 4 & 6 \\ 4 & 2 & 1 \\ 6 & 1 & 6 \end{pmatrix}$$
Since $D_1 = 1$, $D_2 = -14$, and $D_3 = -109$, we get that none of the conditions hold, thus the matrix is indefinite.

2. Consider the matrix
$$A = \begin{pmatrix} 3 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 3 \end{pmatrix}$$
Since $D_1 = 3$, $D_2 = 5$ and $D_3 = 12$ it is positive definite.

3. The symmetric matrix associated to the quadratic form $xy + yz$ is indefinite since $D_1 = D_3 = 0$ and $D_2 = -\frac{1}{4}$.