BACHELOR

Determinantal representation of polynomials

van Doornmalen, M.J.

Award date:
2016

Technische Universiteit Eindhoven

Bachelor Final Project

Determinantal representation of polynomials

Jasper van Doornmalen

supervised by

prof.dr.ir. J. (Jan) Draisma

June 21, 2016

Contents

Chapter 1

Introduction

Many practical problems can be represented by a system of mathematical equations. Because only a small portion of the systems can be solved analytically, numerical methods are required to find approximations of solutions for many types of systems of equations.

In the case of root-finding in a univariate polynomial, there do exist explicit formulas for general polynomials up to degree four [1]. For higher order polynomials, Abel's Impossibility Theorem states that it is impossible to have algebraic solutions in terms of radicals [2]. To find the roots of polynomials of higher degree, one can use an eigenvalue-finding algorithm on a matrix in which the characteristic polynomial is the considered polynomial up to a nonzero scalar factor. This is usually done in numerical software packages [3]. A suitable simple matrix in the univariate case is the companion matrix [4]. When the polynomial is of degree n, then the companion matrix is $n \times n$. Requiring that the entries of the matrix are linear in the variable of the polynomial then this will be a matrix of minimal size.

For many algorithms the complexity is a relatively high-degree polynomial in the dimension of the matrix. For this reason, it is interesting to find matrix representations for polynomials of minimal dimension [3]. This thesis will proceed on the work by Bor Plestenjak and Michiel E. Hochstenbach [3]. Whereas that paper focuses on finding representations for bivariate polynomials, in this thesis the emphasis is targeted at finding representations for any multivariate polynomial of any degree. As in the bivariate representation by Plestenjak and Hochstenbach, this paper also only considers matrices where the entries are affine linear in the coefficients and in the variables of the polynomial.

This report is organized as follows. Chapter 2 lists the notation that is used in the report. In Chapter 3 a number of constructions are considered. In Section 3.1 the general properties of the matrices will be discussed. First the companion matrix will be discussed in Section 3.2. Next the construction by Plestenjak and Hochstenbach will be discussed in Section 3.3, and will be generalized for the multivariate case in Section 3.4. In Section 3.5 a different approach is introduced, which is more efficient when high degrees are considered. In Chapter 4 lower bounds for the matrix dimensions are determined. In Section 4.1 a global lower bound is determined, and in Section 4.2 a few tight lower bounds are determined. Lastly, in conclusion in Chapter 5 the results will be compared and discussed.

Chapter 2

Notation

This section contains a number of notational remarks.

p will be referred to as a polynomial that is of degree d and that has n variables. The variables are $x_1, \ldots, x_n$. This will be denoted as $p \in \mathbb{C}[x_1, \ldots, x_n]_{\leq d}$. The set of variables is referred to as $\mathcal{V}$, and the set of coefficients is referred to as $\mathcal{D}$.

$\mathbb{C}[x_1, \ldots, x_n]_{\leq d}$ is a truncation of the polynomial ring $\mathbb{C}[x_1, \ldots, x_n]$ where all elements of degree larger than d are omitted. Note that $\mathbb{C}[x_1, \ldots, x_n]_{\leq d}$ is not a polynomial ring, because the multiplicative property is not satisfied.

Likewise $\mathbb{C}[x_1, \ldots, x_n]_{=d}$ is defined as every polynomial in $x_1, \ldots, x_n$ in which every term is of degree d.

p can be written in the following two ways:

$$p(x) := \sum_{\alpha \in \mathbb{N}_0^n \,:\, |\alpha| \leq d} c_\alpha x^\alpha \qquad\qquad p(x_1, \ldots, x_n) := \sum_{i_1 + \cdots + i_n \leq d \,:\, i_1, \ldots, i_n \in \mathbb{N}_0} c_{i_1 \ldots i_n} x_1^{i_1} \ldots x_n^{i_n}$$

The sets $\mathcal{V}$ and $\mathcal{D}$ can be written as follows:

$$\mathcal{V} := \{x_1, \ldots, x_n\} \qquad\qquad \mathcal{D} := \{c_\alpha \;:\; \alpha \in \mathbb{N}_0, \alpha \leq d\}$$

M is a square $N \times N$-matrix in which the entries are affine linear in the variables and affine linear in the coefficients: $M \in \mathbb{C}[\mathcal{D}]_{\leq 1}[\mathcal{V}]_{\leq 1}^{N \times N}$. M is often referred to as the matrix that has the polynomial p as determinant.

$$p := p(x_1, \ldots, x_n) = \det M(x_1, \ldots, x_n) =: \det M$$

The matrix M is often written as the sum of two matrices M_0 and M_1: $M = M_0 + M_1$. In this notation M_0 is chosen such that it is independent on the coefficients of p, and M_1 is chosen fully dependent on the coefficients of p:

$$M_0 \in \mathbb{C}[\mathcal{V}]_{\leq 1}^{N \times N} \qquad\qquad M_1 \in \mathbb{C}[\mathcal{D}]_{=1}[\mathcal{V}]_{\leq 1}^{N \times N}$$

Also this notation for spans over some set $\mathbb{F}$ (usually a ring, such as $\mathbb{C}$, $\mathbb{R}$ or $\mathbb{Z}$) is used:

$$\langle P_1, \ldots, P_n \rangle_{\mathbb{F}} = \{\lambda_1 P_1 + \cdots + \lambda_n P_n : \lambda_1, \ldots, \lambda_n \in \mathbb{F}\}$$

Lastly, $\mathbb{N}_k := \{k, k+1, k+2, \ldots\}$ is defined as the set of nonnegative integers added to k. For this reason, $\mathbb{N}_0$ is the set of nonnegative integers, and $\mathbb{N}_1$ is the set of positive integers.

Chapter 3

Constructions

This chapter focuses on finding general constructions of matrices for polynomials of any degree and any number of variables.

3.1 Properties of the matrix

Let $p \in \mathbb{C}[x_1, \ldots, x_n]_{\leq d}$ be a polynomial with coefficients c_α. Let $M \in \mathbb{C}[\mathcal{D}]_{\leq 1}[\mathcal{V}]_{\leq 1}^{N \times N}$ such that $\det M = p$ for all values of the variables. Use the split $M = M_0 + M_1$ where $M_0 \in \mathbb{C}[\mathcal{V}]_{\leq 1}^{N \times N}$ and $M_1 \in \mathbb{C}[\mathcal{D}]_{=1}[\mathcal{V}]_{\leq 1}^{N \times N}$.

Proposition 3.1.1 *For all $x_1, \ldots, x_n \in \mathbb{C}$ it holds that $\det M_0(x_1, \ldots, x_n) = 0$.*

Proof M_0 is independent of the choice of the coefficients of the polynomial p. Therefore one can choose p without loss of generality for statements about M_0.

Let p be the zero polynomial: $p(x_1, \ldots, x_n) = 0$ for all $x_1, \ldots, x_n$.

The entries of M_1 are from $\operatorname{span}(\mathcal{D} \cdot \mathcal{V} \cup \mathcal{D})$. Because all coefficients of p are zero, $\mathcal{D} = \{0\}$. Therefore all entries from M_1 are zero, so M_1 is the zero matrix: $M_1 = O$.

$$0 = p = \det M = \det(M_0 + M_1) = \det(M_0 + O) = \det M_0$$

Therefore it is found that $\det M_0 = 0$ for all $x_1, \ldots, x_n \in \mathbb{C}$. □

Proposition 3.1.2 *For all $x_1, \ldots, x_n \in \mathbb{C}$ it holds that $\det M_1(x_1, \ldots, x_n) = 0$.*

Proof Suppose $\det M_1 \neq 0$.

Then N nonzero entries of M_1 get multiplied. The entries of M_1 are from $\operatorname{span}(\mathcal{V} \cdot \mathcal{D} \cup \mathcal{D})$, so multiplying N entries results in a polynomial of which every term is a monomial in the variables $\mathcal{V}$ of maximal degree N. The coefficients of these monomials are from $\{\prod_{i=1}^N d_i | d_1, \ldots, d_N \in \mathcal{D}\}$. For a general set of coefficients $\mathcal{D}$, this is not a subset of $\mathcal{D}$.

$$\left\{ \prod_{i=1}^{N} d_i | d_1, \ldots, d_N \in \mathcal{D} \right\} \not\subseteq \mathcal{D}$$

Because all terms that occur in $\det M_1$ do also occur in $\det M$, the polynomial p must contain terms of which the coefficient is the product of N coefficients of itself. For a general polynomial, this does not hold $\lightning$. This leads to a contradiction.

Therefore, $\det M_1 = 0$ for all $x_1, \ldots, x_n \in \mathbb{C}$. $\square$

Proposition 3.1.3 *For all* $x_1, \ldots, x_n$ *it holds that* $\operatorname{rank} M_0 = N - 1$.

Proof From proposition 3.1.1 it follows that M_0 is a nonsingular matrix. From this it follows that $\operatorname{rank} M_0 \leq N - 1$. **(Result 1)**

Suppose there are $x_1, \ldots, x_n \in \mathbb{C}$ such that $\operatorname{rank} M_0(x_1, \ldots, x_n) < N - 1$. Because M_0 is independent of p, without loss of generality one can choose p such that $p(x_1, \ldots, x_n) \neq 0$.

Denote $\underline{x} = (x_1, \ldots, x_n)$. Because $\operatorname{rank} M_0(\underline{x}) < N - 1$, there exists a transformation matrix $T \in \mathcal{SL}_N(\mathbb{C})$ such that $M_0(\underline{x})T = (\underline{b_1}, \ldots, \underline{b_{N-2}}, \underline{0}, \underline{0})$.

Because $p(\underline{x}) \neq 0$, $\det M$ must be nonzero, and so must $\det MT$ be nonzero.

Compute $\det MT$ by expanding the last two columns into minors. Because $M_0(\underline{x})T$ contains zero values in these columns, the entries in the last two columns are the entries from $M_1(\underline{x})T$, so they are element of $\operatorname{span}(\mathcal{D} \cdot \mathcal{D})$.

$$0 \neq \underbrace{p(\underline{x})}_{\in \operatorname{span}(\mathcal{D})} = \det M(\underline{x}) = \underbrace{\det T}_{=1} \det M(\underline{x}) = \underbrace{\det(M_0(\underline{x})T + M_1(\underline{x})T)}_{\in \operatorname{span}(\mathcal{D} \cdot \mathcal{D})}$$

What follows is that $p(\underline{x}) \in \operatorname{span}(\mathcal{D})$ and $p(\underline{x}) \in \operatorname{span}(\mathcal{D} \cdot \mathcal{D})$ $\lightning$. This is not possible.

So for all $x_1, \ldots, x_n$ it holds that $\operatorname{rank} M_1(x_1, \ldots, x_n) \geq N - 1$. **(Result 2)**

Combine Result 1 and Result 2 to find that for all $x_1, \ldots, x_n \in \mathbb{C}$ $\operatorname{rank} M_0 = N - 1$. $\square$

3.2 Frobenius Companion Matrix

Let p be a complex polynomial of degree d with variable x: $p \in \mathbb{C}[x]_{\leq d}$. The coefficients of the polynomial is c_i, where i is the index referring to the degree of the monomial.

$$p(x) = \sum_{i=0}^{d} c_i x^i = c_0 + c_1 x + \cdots + c_d x^d$$

In this section a matrix representation for univariate polynomials is discussed. Requiring that the matrix is linear in the variables and coefficients of the polynomial, a lower bound for the matrix dimension is the degree of the polynomial, d.

The Frobenius companion matrix is a well-known matrix representation of univariate polynomials. The characteristic polynomial of the Frobenius companion matrix will be a monic univariate polynomial. It holds that $p(x) = \det(xI - M)$. Let C be the Frobenius Companion matrix that is defined as follows [4]:

$$C = \begin{pmatrix} 0 & \dots & \dots & 0 & -c_0 \\ 1 & \ddots & \ddots & 0 & -c_1 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & & \ddots & 0 & -c_{d-2} \\ 0 & \dots & & 1 & -c_{d-1} \end{pmatrix}$$

If C is the companion matrix as above, then the characteristic polynomial is the monic polynomial $p(x) = \sum_{i=0}^{d} c_i x^i = c_0 + c_1 x + \cdots + c_d x^d$. Because it is monic, the coefficient for x^d will be one: $c_d = 1$. The Frobenius companion matrix determinantal representation of p follows by computing the characteristic polynomial:

$$p(x) = \det(xI - M) = \det \begin{pmatrix} x & \dots & \dots & 0 & c_0 \\ -1 & \ddots & \ddots & 0 & c_1 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & & \ddots & x & c_{d-2} \\ 0 & \dots & & -1 & x + c_{d-1} \end{pmatrix}$$

The dimension of this matrix is $d \times d$, and the entries are linear in the coefficients and in the variables. With a lower dimension it's not possible to get a term containing the monomial x^d. For this reason, the dimension $d \times d$ is minimal for the univariate case.

3.3 Construction for bivariate polynomials by Plestenjak and Hochstenbach

For polynomials in two variables a simple method is known to construct a matrix that is affine linear in the coefficients and affine linear in the variables of the polynomial. This section will discuss this construction, which is proposed by the paper from Bor Plestenjak and Michiel E. Hochstenbach [3].

Let $q \in \mathbb{C}[x, y]$ be some bivariate polynomial, and let T be the set of monomials occurring in this polynomial. Let set S be such that $T \subseteq \{1, x, y\} \cdot S$, and that for all $m \in S$ either $m = 1$ or there is another monomial $\widetilde{m} \in S$ such that $m = x\widetilde{m}$ or $m = y\widetilde{m}$.

$$S = \{m \mid m = 1 \vee \exists \widetilde{m} \in S : m \in \{x, y\} \cdot \widetilde{m}\}$$

This can be represented by a directed rooted acyclic graph representation. Let the root of the graph be the monomial 1 and the left child of the graph correspond to the monomial x multiplied with its parent, and the right child of the graph to the monomial y multiplied with its parent. This is visualized in Figure 3.1. The set S will contain the parental nodes of the nodes on the ends of the edges that are needed to connect the root node with the monomials of T. If S is chosen to be minimal, the resulting graph that only contains nodes from S is a directed tree.

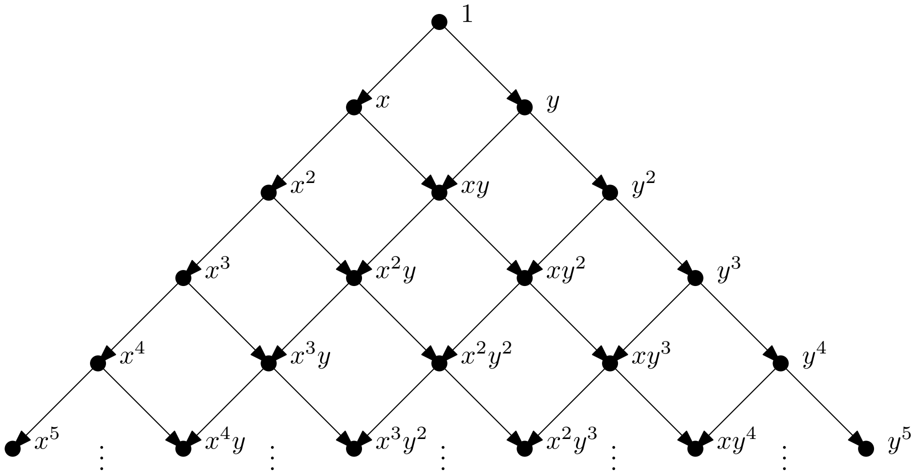


Figure 3.1: Monomial graph

Write $S = \{1 = m_1, m_2, \ldots, m_N\}$ such that $\forall i : \deg(m_i) \leq \deg(m_{i+1})$. Also write $m_i = x_{j_i} \cdot m_{a_i}$ where $x_{j_i} \in \{1, x, y\}$. Since $\forall i : \deg(m_i) \leq \deg(m_{i+1})$ it holds that $a_i < i$ for all $i > 1$.

Let $M_S := M_S(x, y)$ be a $N \times N$ matrix. Let $\underline{e_i}$ be the i'th standard basis row vector of length N. For the rows 1 to $N-1$, define the row $i-1$ as follows:

$$x_{j_i}\underline{e_{a_i}} - \underline{e_i} = (0, \ldots, x_{j_i}, \ldots, -1, 0, \ldots, 0)$$

And define row N as vector $\underline{r}$ such that the matrix product (analogous to the standard inner product) $\underline{r} \cdot (1, m_2, m_3, \ldots, m_N)^T = q$. It is proposed that $\det M_S(x, y) = q(x, y)$, of which a more general case will be proved in Section 3.4.1.

The entries of $\underline{r}$ should be affine linear combinations of the coefficients of the polynomial q, and of the variables $x_1, \ldots, x_n$ and 1. This is possible, because every term from the polynomial q could be created by multiplying the coefficient of the term, with a monomial m_i and with one element out of $\{1, x_1, \ldots, x_n\}$. Because all monomials m_i are in the vector that $\underline{r}$ is multiplied with, only the products between the coefficient and an element from $\{1, x_1, \ldots, x_n\}$ need to occur in $\underline{r}$.

The vector $\underline{r}$ is not necessarily unique, because if there exist two different monomials m_i and m_j such that $\exists x_k, x_l : x_k m_i = x_l m_j$, then the term with monomial $x_k m_i$ could be either part of entry r_i, or be part of entry r_j. Therefore, these could be interchanged.

By choosing the polynomial q to be a general bivariate polynomial of a certain total degree, this construction will offer a matrix that has this polynomial as its determinant.

3.3.1 Example

The construction will be shown with an example. Let p be a general bivariate polynomial of degree 3.

$$p(x, y) = x^3 c_{3,0} + x^2 y c_{2,1} + x^2 c_{2,0} + x y^2 c_{1,2} + x y c_{1,1} + x c_{1,0} + y^3 c_{0,3} + y^2 c_{0,2} + y c_{0,1} + c_{0,0}$$

The set T contains all monomials, so $T = \{1, y, y^2, y^3, x, xy, xy^2, x^2, x^2y, x^3\}$.

S can be selected as you like, but it must contain 1, must be connected and $\{1, x, y\} \cdot S \supseteq T$. The choices $S = \{1, x, y, x^2, y^2\}$ are valid, because it satisfies these conditions. This is shown in Figure 3.2. Also keep the monomials of set S in the given order: $(m_1, m_2, m_3, m_4, m_5) = (1, x, y, x^2, y^2)$. Because $|S| = 5$ the matrix M_S will be 5×5.

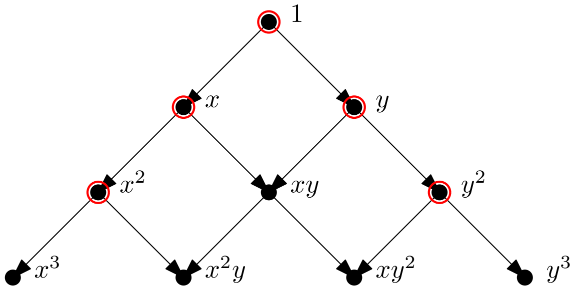


Figure 3.2: Monomial graph showing the choices of S in example.
Note that every non-selected point is neighbor of a selected point.

First consider row i with $i - 1 = 1$, so $i = 2$. $m_i = m_2 = x_{j_2} \cdot m_{a_2} = x$, so $a_2 = 1$ and $x_{j_2} = x$. Therefore, the first row will be $x\underline{e_1} - \underline{e_2} = (x, -1, 0, 0, 0)$.

Doing the same for $i = 3, i = 4$ and $i = 5$ to find the rows 2, 3 and 4. Row 2 will be $(y, 0, -1, 0, 0)$, row 3 will be $(0, x, 0, -1, 0)$ and row 4 will be $(0, 0, y, 0, -1)$.

Re-writing polynomial p shows how the last column can be chosen. There are multiple ways to write this polynomial, in this example the following way is chosen:

$$p(x, y) = 1\cdot(c_{0,0}+c_{1,0}x+c_{0,1}y)+x\cdot(c_{2,0}x+c_{1,1}y)+y\cdot(c_{0,2}y)+x^2\cdot(c_{3,0}x+c_{2,1}y)+y^2\cdot(c_{1,2}x+c_{0,3}y)$$

From this it follows that the last row is:

$$(c_{0,0} + c_{1,0}x + c_{0,1}y,\ c_{2,0}x + c_{1,1}y,\ c_{0,2}y,\ c_{3,0}x + c_{2,1}y,\ c_{1,2}x + c_{0,3}y)$$

The result is the matrix $M_S(x, y)$:

$$M_S(x, y) = \begin{pmatrix} x & -1 & 0 & 0 & 0 \\ y & 0 & -1 & 0 & 0 \\ 0 & x & 0 & -1 & 0 \\ 0 & 0 & y & 0 & -1 \\ c_{0,0} + yc_{0,1} + xc_{1,0} & yc_{1,1} + xc_{2,0} & yc_{0,2} & yc_{2,1} + xc_{3,0} & yc_{0,3} + xc_{1,2} \end{pmatrix}$$

It is easily verified that $\det M_S(x, y) = p(x, y)$.

3.4 Generalized method from Plestenjak and Hochstenbach

The proposed method of the previous section can be generalized for higher degrees.

Let p be some general polynomial with n variables and of total degree d. Use $\mathcal{V} = \{x_1, \ldots, x_n\}$ as the set of variables, and $\mathcal{D}$ the set of coefficients, then we can write $p \in \mathbb{C}[\mathcal{V}]_{\leq d}$. As in the previous section, let T be the set of all monomials in p:

$$T = \{\text{Monomials of polynomial } p\} = \{x_1^{i_1} x_2^{i_2} \ldots x_n^{i_n} \mid i_1, \ldots, i_n \in \mathbb{N}_0,\ i_1 + i_2 + \cdots + i_n \leq d\}$$

More specifically, if p is a more specific polynomial with fewer monomials, then the set T can be adjusted as such. This will not break the construction, and may lead to a matrix of smaller dimension.

Furthermore, let S be a set of monomials such that $\forall m \in T\ \exists \widetilde{m} \in S\ :\ m \in \widetilde{m} \cdot \{1, x_1, \ldots, x_n\}$, and where every monomial of S that is not 1 has a parent in S. This is analogous to set S in the previous section for more variables.

$$S = \{m | m = 1 \vee \exists \widetilde{m} \in S\ :\ m \in \widetilde{m} \cdot \mathcal{V}\}$$

Analogous to the bivariate case, write $S = \{1 = m_1, m_2, \ldots, m_N\}$ such that $\forall i\ :\ \deg(m_i) \leq \deg(m_{i+1})$, and write $m_i = x_{j_i} \cdot m_{a_i}$, where $x_{j_i} \in \{1\} \cup \mathcal{V} = \{1, x_1, \ldots, x_n\}$. Also, $a_i < i$ for all $i > 1$ because for all i it holds that $\deg(m_i) \leq \deg(m_{i+1})$.

The construction of the matrix of which the determinant is the polynomials is analogous to the bivariate case. Let $M_S := M_S(x, y)$ be a $N \times N$-matrix, and let $\underline{e_i}$ be the i'th basic row vector of length $N = |S|$. For the rows 1 to $N - 1$, define row $i - 1$ as follows:

$$x_{j_i} \underline{e_{a_i}} - \underline{e_i} = (0, \ldots, x_{j_i}, \ldots, -1, 0, \ldots, 0)$$

And let row N be a vector $\underline{r}$ such that the matrix product $\underline{r} \cdot (1, m_2, m_3, \ldots, m_N)^T = p(x_1, \ldots, x_n)$. The components of the row vector $\underline{r}$ can be chosen affine linear in the variables and in the coefficients of the polynomial p, because every monomial from p can be written as the product of one of the variables with a monomial $1, m_2, m_3, \ldots, m_N$.

The shape of the matrix follows:

$$M_S(x_1, \ldots, x_n) = \begin{pmatrix}
m_1 & -1 & 0 & & & & \cdots & & & 0 \\
m_2 & 0 & -1 & 0 & & & \cdots & & & 0 \\
m_3 & 0 & 0 & -1 & 0 & & \cdots & & & 0 \\
\vdots & & & & \ddots & & & & & 0 \\
m_n & 0 & 0 & & \cdots & -1 & 0 & & \cdots & 0 \\
0 & * & * & & \cdots & & -1 & 0 & \cdots & 0 \\
 & \vdots & & \vdots & & & & \ddots & \vdots & \vdots \\
* & & \cdots & & \cdots & & \cdots & * & -1 & \\
* & & \cdots & & \cdots & & \cdots & * & * & -1 \\
\underline{r_1} & \underline{r_2} & \underline{r_3} & & \cdots & & \cdots & & & \underline{r_N}
\end{pmatrix}$$

And it's proposed that $\det M_S(x_1, \ldots, x_n) = p(x_1, \ldots, x_n)$, and this will be proved next.

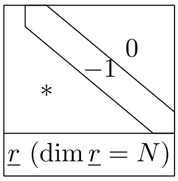
3.4.1 Proof of correctness

This is a proof by induction to the cardinality of the set S, which is the dimension of the matrix M_S: $|S| = N$. For simplicity, write $\underline{m_k} = (1 = m_1, m_2, \ldots, m_k)$ as row vector of the ordered monomials of S. The proof will use the property that the inner product of the bottom row of the matrix with the ordered monomial vector is the polynomial: $p = \underline{r} \cdot \underline{m_n}$.

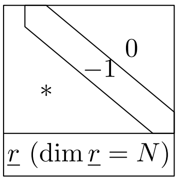
Induction Basis: Let $N = 1$, then there's only one monomial $m_1 = 1 \in S$. Then $M_S = (r_1)$ is the 1×1 matrix and indeed we find that $\det(M_S) = r_1 = r_1 m_1 = \underline{r} \cdot \underline{m_1} = p$. So the proposition holds for $|S| = N = 1$.

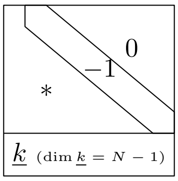
Induction Hypothesis: Suppose that $\det(M_S) = p$ for all $(N-1) \times (N-1)$ matrices M_S (that satisfy the construction) with $|S| \leq N - 1$.

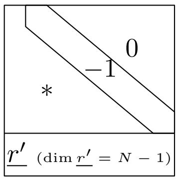
Induction Step: Consider matrix M_S of size $N \times N$ where $|S| = N$. Here, M_S is the constructed matrix of a general polynomial p, so that $p = \underline{r} \cdot \underline{m_N}$. A visualization of the matrix follows:

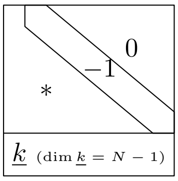
$M_S =$ 

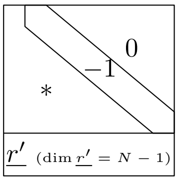
In this matrix, we compute the determinant by expanding the last column into minors. Note that the last column of the matrix M_S is $(0, \ldots, 0, -1, r_N)^T$, so the expansion leads to two terms only. We find the following result:

$\det M_S = \det$ 

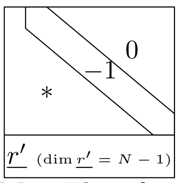
$= (-1)^{2N} r_N \det$ 

$+ (-1)^{2N-1}(-1) \det$ 

$= r_N \det$ 

$+ \det$ 

In this notation, $\underline{r'} = (r_1, \ldots, r_{N-1})$, and the vector $\underline{k}$ is such that $\underline{k} \cdot \underline{m_{N-1}} - m_N = 0$ and only one entry from the vector $\underline{k}$ is nonzero. This follows from the structural definition of the matrix M_S. It holds that $\underline{k} = x_i \underline{e_j}$ for certain i and j.

The determinant from the second term – $\det$ 

– is a determinant of a $(N-1) \times (N-1)$ matrix with the same structure as M_S. Therefore, by the induction hypothesis the determinant of this matrix is as follows:

$$\underline{r'} \cdot \underline{m_{N-1}} = (r_1, \ldots, r_{N-1}) \cdot (m_1, \ldots, m_{N-1}) = \sum_{i=1}^{N-1} r_i m_i$$

The determinant of the first term – $\det \left|\begin{array}{c} \begin{matrix} & & 0 \\ & -1 & \\ * & & \end{matrix} \\ \hline \underline{k} \ (\dim \underline{k} = N-1) \end{array}\right|$ – is also a determinant of a $(N-1) \times (N-1)$ matrix that has the same structural form of matrix M_S. Since this is a $(N-1) \times (N-1)$ matrix, it holds that this determinant is the same as $\underline{k} \cdot \underline{m_{N-1}}$. From the original matrix, it's known that $\underline{k} \cdot \underline{m_{N-1}} - m_N = 0$, so it follows that $\underline{k} \cdot \underline{m_{N-1}} = m_N$. Therefore $\det \left|\begin{array}{c} \begin{matrix} & & 0 \\ & -1 & \\ * & & \end{matrix} \\ \hline \underline{k} \ (\dim \underline{k} = N-1) \end{array}\right| = m_N$.

By using these two results in the derived equality for $\det M_S$ yields:

$$\det M_S = \sum_{i=1}^{N-1} (r_i m_i) + r_N m_N = \sum_{i=1}^{N} (r_i m_i) = \underline{r} \cdot \underline{m_N} = p$$

So for a $|S| = N$ it also holds that $\det M_S = p$. So by induction it follows that for all $|S| = N \geq 1$ holds that $\det M_S = p$. $\square$

3.4.2 Asymptotic analysis

Because the dimension of matrix M equals the cardinality of the set S, the matrix dimensions in this construction are minimal when the set S is chosen to be a minimal set.

The paper by Plestenjak and Hochstenbach discusses that for finding such a minimal set S for any polynomial corresponds to a NP-hard directed Steiner tree problem. However, they do note that there exist algorithms running in polynomial time that gives approximate solutions that are close to the minimal solution [3].

The paper furthermore contains a choice for the set S in the bivariate case for all possible degrees. However, it is possible that there is a better construction, since minimality is not guaranteed [3]. Next the closed-form formula for the cardinality of S will be derived for that construction. Then this method will be generalized for more variables, and the asymptotic cardinality will be determined.

Denote $\varphi(d)$ as the number of monomials in the set S for bivariate polynomials of degree d, using the construction of Plestenjak and Hochstenbach. If the degree is 1, then $S = \{1\}$ is minimal, so $\varphi(1) = 1$. If the degree is 2, then $S = \{1, x, y\}$ is minimal, so $\varphi(2) = 3$. For higher degrees, the following construction can be used. Let S_d be the minimal set of monomials for degree d, then $S_{d+2} = \{1, y,\ x, x^2, \ldots, x^{n-1}\} \cup y^2 \cdot S_d$. Obviously, the sets $\{1, y,\ x, x^2, \ldots, x^{n-1}\}$

and $y^2 \cdot S_d$ are disjoint. Therefore, $\varphi(d+2) = |S_{d+1}| = |\{1, y,\ x, x^2, \ldots, x^{n-1}\} \cup y^2 \cdot S_d| = |\{1, y,\ x, x^2, \ldots, x^{n-1}\}| + |y^2 \cdot S_d| = (n+1) + \varphi(d)$.

Using this recursion, one can find the explicit form.

$$\left.\begin{aligned} \varphi(d) = d + 1 + \varphi(d-2) \\ \varphi(1) = 1 \\ \varphi(2) = 3 \end{aligned}\right\} \Rightarrow \varphi(d) = \begin{cases} \frac{1}{4}d(d+4) & d \text{ even} \\ \frac{1}{4}(d^2+4d-1) & d \text{ odd} \end{cases}$$ [1]

So in the bivariate case the construction is asymptotically of order $\frac{1}{4}d^2$.

For a higher number of variables such a choice for the set S is not known. However, it's possible to generalize the method to find the set S in the bivariate case for a higher number of variables. Next will be given a construction for the set S with any number of variables, but it should be noted that there is no statement about minimality of this construction.

Consider a polynomial p of degree d in n variables. Denote $S_{n,d}$ as the set of chosen monomials in the case of degree d and n variables. Denote $\varphi(n,d)$ by the cardinality of the set $S_{n,d}$: $\varphi(n,d) = |S_{n,d}|$. Let the variables be $x_1, x_2, \ldots, x_n$.

When the degree is 1, then $S_{n,1} = \{1\}$ is minimal, so $\varphi(n,1) = 1$. When the degree is 2, then $S_{n,2} = \{1, x_1, x_2, \ldots, x_n\}$ is minimal, so $\varphi(n,2) = n+1$. When there is only one variable, the set $S_{1,d} = \{1, x_1, x_1^2, \ldots, x_1^{n-1}\}$ is minimal, so $\varphi(1,d) = d$. As in the bivariate case, the set $S_{n,d}$ will be constructed by adding monomials to the set $x_n^2 \cdot S_{n,d-2}$. The formal definition of the generalized construction of $S_{n,d}$ follows:

$$S_{n,d} = x_n^2 \cdot S_{n,d-2} \cup \{x_n\} \cup \{x_1^{i_1} \ldots x_{n-1}^{i_{n-1}} | i_1 + \cdots + i_{n-1} \leq d-1, i_1, \ldots, i_{n-1} \in \mathbb{N}_0\} \quad (3.1)$$

It is easily verified that indeed every monomial up to degree d can be created from an entry from this set multiplied with a monomial from $\{x_1, \ldots, x_n\}$. If the monomial contains a factor x_n^2 then this can be found by multiplying an element from $x_n^2 \cdot S_{n,d-2}$ with a variable, except if the monomial equals x_n^2, but this can be found from the added monomial x_n. If the monomial does not contain a factor x_n^2, it can be found by multiplying an element from $\{x_1^{i_1} \ldots x_{n-1}^{i_{n-1}} | i_1 + \cdots + i_{n-1} \leq d-1, i_1, \ldots, i_{n-1} \in \mathbb{N}_0\}$ with a variable.

Now compute $\varphi(n,d)$:

$$\begin{aligned} \varphi(n,d) &= |S_{n,d}| \\ &= \left| x_n^2 \cdot S_{n-2,d} \cup \{x_n\} \cup \{x_1^{i_1} \ldots x_{n-1}^{i_{n-1}} | i_1 + \cdots + i_{n-1} \leq d-1, i_1, \ldots, i_{n-1} \in \mathbb{N}_0\} \right| \\ &= \varphi(n, d-2) + 1 + \left| \{x_1^{i_1} \ldots x_{n-1}^{i_{n-1}} | i_1 + \cdots + i_{n-1} \leq d-1, i_1, \ldots, i_{n-1} \in \mathbb{N}_0\} \right| \\ &\overset{(*)}{=} \varphi(n, d-2) + 1 + \binom{(n-1)+(d-1)}{d-1} \end{aligned}$$

[1] Note that this closed-form result is different than the result of page 9 in the paper of Plestenjak and Hochstenbach [3]. I have good belief in the correctness of this derivation, because this result matches the given part of their sequence, and because their result returns fractional values for φ. For this reason, the closed-form derivation of their paper cannot be correct.

The last part of this equation is found by evaluating the cardinality of the set $\{x_1^{i_1} \dots x_{n-1}^{i_{n-1}} | i_1 + \dots + i_{n-1} \leq d-1, i_1, \dots, i_{n-1} \in \mathbb{N}_0\}$, which is the number of solutions of the following system:

$$\begin{cases} i_1 + i_2 + \dots + i_{n-1} \leq d - 1 \\ i_1, i_2, \dots, i_{n-1} \in \mathbb{N}_0 \end{cases}$$

By introducing a dummy variable $\eta = d - (i_1 + i_2 + \dots + i_n) \in \mathbb{N}_0$, the problem can be re-written as

$$\begin{cases} i_1 + i_2 + \dots + i_{n-1} + \eta = d - 1 \\ i_1, \dots, i_{n-1}, \eta \in \mathbb{N}_0 \end{cases}$$

This is a known problem and the number of solutions is $\binom{(n-1+1)+(d-1)-1}{d-1} = \binom{n+d-2}{d-1}$ [5].

The recurrence equation that's found for higher dimensions is:

$$\begin{cases} \varphi(n, d) = \varphi(n, d-2) + 1 + \binom{(n-1)+(d-1)}{d-1} \\ \varphi(n, 1) = 1 \\ \varphi(n, 2) = n + 1 \end{cases}$$

In Table 3.1 closed-form polynomial expressions of $\varphi(n, d)$ are given as a function of d with a fixed number of variables n. Because for two variables the generalized construction is the same as the construction of the bivariate case, the values of $\varphi(2, d)$ are the same as the bivariate case.

n	d even	d odd
1	d	d
2	$\frac{d^2}{4} + d$	$\frac{d^2}{4} + d - \frac{1}{4}$
3	$\frac{d^3}{12} + \frac{3d^2}{8} + \frac{11d}{12}$	$\frac{d^3}{12} + \frac{3d^2}{8} + \frac{11d}{12} - \frac{3}{8}$
4	$\frac{d^4}{48} + \frac{d^3}{6} + \frac{5d^2}{12} + \frac{5d}{6}$	$\frac{d^4}{48} + \frac{d^3}{6} + \frac{5d^2}{12} + \frac{5d}{6} - \frac{7}{16}$

Table 3.1: Table of $\varphi(n, d)$ for a specific number of variables n.

An asymptotic bound for $\varphi(n, d) = |S_{n,d}|$ will be derived using a geometric approach. Represent the set $S_{n,d}$ by a n-dimensional vector space which is subset of $\mathbb{N}_0^n$. Denote $S_{n,d}$ in lattice point representation as $\hat{S}_{n,d}$. Each monomial $x_1^{i_1} \dots x_n^{i_n}$ will be represented by the lattice points $(i_1, \dots, i_n)$. The recursive Equation 3.1 considers all monomials in the set $S_{n,d-2}$, multiplies these with x_n^2 and adds the monomials x_n and $x_1^{i_1} \dots x_{n-1}^{i_{n-1}}$ for all $i_1 + \dots + i_{n-1} \leq d - 1$. In lattice point representation, the set $\hat{S}_{n,d}$ is recursively defined as follows:

$$\hat{S}_{n,d} = \{(i_1, \dots, i_{n-1}, 0) \mid i_1 + \dots + i_{n-1} \leq d - 1\} \cup \{(0, \dots, 0, 1)\} \cup \Big(\hat{S}_{n,d-2} + (0, \dots, 0, 2)\Big)$$

The construction can be described as follows. To construct a set $S_{n,d}$, use the already known set $S_{n,d-2}$ and add 2 to every last lattice point entry. Add the point $(0, \dots, 0, 1)$ to this set, and all lattice points that has the last lattice point entry 0.

Denote the set $\hat{U}_{n,d}$ as the set containing all monomials of total degree smaller than $d-1$ in lattice point representation. This is defined as follows:

$$\hat{U}_{n,d} = \{(i_1, i_2, \ldots, i_n) \in \mathbb{N}_0^n \mid i_1 + i_2 + \cdots + i_n \leq d-1\}$$

Note that the cardinality of this set is $|\hat{U}_{n,d}| = \binom{n+(d-1)}{(d-1)} = \frac{\prod_{l=1}^{n}((d-1)+l)}{n!}$, so the leading term of $|\hat{U}_{n,d}|$ is $\frac{d^n}{n!}$. More formally this result can be written as a limit:

$$\lim_{d\to\infty} \frac{|\hat{U}_{n,d}|}{\frac{d^n}{n!}} = 1$$

Using these notations, the set $\hat{S}_{n,d}$ can be written as a set without recursion involved. The set $\hat{S}_{n,d}$ contains all lattice points of $\hat{U}_{n,d}$ where $i_n \equiv 0 \mod 2$, along with the lattice points $(0, \ldots, 0, i_n)$ where $i_n \equiv d-1 \mod 2$. More formally this can be written as:

$$\hat{S}_{n,d} = \{(i_1, \ldots, i_n) \in \hat{U}_{n,d} \mid i_n \equiv 0 \mod 2\} \cup \{(0, \ldots, 0, i_n) \in \hat{U}_{n,d} \mid i_n \equiv 1 \mod 2\}$$

It is easily verified that $\hat{S}_{n,d} \cup \Big((0, \ldots, 0, 1) + \hat{S}_{n,d-1}\Big) = \hat{U}_{n,d}$, because $\hat{S}_{n,d}$ contains all points with $i_n \equiv 0 \mod 2$ and of maximal degree $d-1$, and $(0, \ldots, 0, 1) + \hat{S}_{n,d-1}$ contains all points with $i_n \equiv 1 \mod 2$ and of maximal degree $d-1$. So together they contain all points in $\hat{U}_{n,d}$.

Using this result, the cardinality of the set $\hat{U}_{n,d}$ can be written as follows:

$$|\hat{U}_{n,d}| = \Big|\hat{S}_{n,d}\Big| + \Big|(0, \ldots, 0, 1) + \hat{S}_{n,d-1}\Big| - \Big|\hat{S}_{n,d} \cap \Big((0, \ldots, 0, 1) + \hat{S}_{n,d-1}\Big)\Big|$$

It holds that $\Big|\hat{S}_{n,d} \cap \Big((0, \ldots, 0, 1) + \hat{S}_{n,d-1}\Big)\Big| = d-1$, because this set corresponds to all points $(0, \ldots, 0, i_n) \in \hat{U}_{n,d}$ with $1 \leq i_n \leq d-1$.

Next will be shown that the cardinalities of $\hat{S}_{n,d}$ and $\hat{S}_{n,d-1}$ have a leading term that is asymptotically the same.

Denote $\hat{T}_{n,d} = \{(i_1, \ldots, i_n) \in \hat{U}_{n,d} \mid i_n \equiv 0 \mod 2\}$. It will follow that the cardinality of this set defines the asymptotic behavior of $|\hat{S}_{n,d}|$.

It holds that the set $\hat{S}_{n,d} = \hat{T}_{n,d} \cup \{(0, \ldots, 0, i_n) \in \hat{U}_{n,d} \mid i_n \equiv 1 \mod 2\}$, from which it follows that $|\hat{S}_{n,d}| = |\hat{T}_{n,d}| + \lfloor \frac{d}{2} \rfloor$. For all n and d it holds that $\hat{T}_{n,d-1} \subseteq \hat{T}_{n,d}$ as is shown next:

$$\begin{aligned}
\hat{T}_{n,d-1} &= \{(i_1, \ldots, i_n) \in \hat{U}_{n,d-1} \mid i_n \equiv 0 \mod 2\} \\
&= \{(i_1, \ldots, i_n) \in \mathbb{N}_0^n \mid i_1, \ldots, i_n \leq d-2 \ , \ i_n \equiv 0 \mod 2\} \\
&\subseteq \{(i_1, \ldots, i_n) \in \mathbb{N}_0^n \mid i_1, \ldots, i_n \leq d-1 \ , \ i_n \equiv 0 \mod 2\} \\
&= \{(i_1, \ldots, i_n) \in \hat{U}_{n,d} \mid i_n \equiv 0 \mod 2\} = \hat{T}_{n,d}
\end{aligned}$$

This result yields that $|\hat{T}_{n,d}| \geq |\hat{T}_{n,d-1}|$. Now combine the derived results above:

$$\varphi(n,d) = |S_{n,d}| = |\hat{S}_{n,d}| = |\hat{T}_{n,d}| + \left\lfloor \frac{d}{2} \right\rfloor \geq |\hat{T}_{n,d-1}| + \left\lfloor \frac{d-1}{2} \right\rfloor = |\hat{S}_{n,d-1}| = |S_{n,d-1}| = \varphi(n, d-1)$$

What follows is that $|\hat{S}_{n,d}| \geq |\hat{S}_{n,d-1}| \geq \cdots \geq |\hat{S}_{n,1}| \geq 0$, which corresponds by using the original notation with $\varphi(n,d) \geq \varphi(n,d-1) \geq \cdots \geq \varphi(n,1) \geq 0$.

Now use the result that $|\hat{U}_{n,d}| = |\hat{S}_{n,d}| + |\hat{S}_{n,d-1}| - |\hat{S}_{n,d} \cap \hat{S}_{n,d-1}| = |\hat{S}_{n,d}| + |\hat{S}_{n,d-1}| - d + 1$. It holds that $\hat{U}_{n,d}$ resembles $\frac{d^n}{n!}$ as d is large, due to $\lim\limits_{d\to\infty} \frac{|\hat{U}_{n,d}|}{\frac{d^n}{n!}} = 1$.

$$1 = \lim_{d\to\infty} \frac{|\hat{U}_{n,d}|}{\frac{d^n}{n!}} = \lim_{d\to\infty} \frac{|\hat{S}_{n,d}| + |\hat{S}_{n,d-1}| - d + 1}{\frac{d^n}{n!}} = \lim_{d\to\infty} \left(\frac{|\hat{S}_{n,d}|}{\frac{d^n}{n!}} + \frac{|\hat{S}_{n,d-1}|}{\frac{d^n}{n!}} + \frac{-d+1}{\frac{d^n}{n!}} \right)$$

If $n \geq 2$, then the last term has no contribution to the limit. For this reason, the case $n = 1$ will not be considered in this computation, and this will not lead to major problems. The case $n = 1$ corresponds to the companion matrix, which is minimal already: the matrix size is exactly $d \times d$. Therefore analyzing the case $n = 1$ is not necessary. So only consider the case $n \geq 2$, such that the last term will not give a contribution in the limit.

Since $|\hat{S}_{n,d}| \geq 0$, it follows that $|\hat{S}_{n,d}|$ and $|\hat{S}_{n,d-1}|$ are upper bounded by $|\hat{U}_{n,d}|$ which is asymptotically $\frac{d^n}{n!}$, and lower bounded by zero. Therefore the limits $\lim\limits_{d\to\infty} \frac{|\hat{S}_{n,d}|}{\frac{d^n}{n!}}$ and $\lim\limits_{d\to\infty} \frac{|\hat{S}_{n,d-1}|}{\frac{d^n}{n!}}$ will exist, so the limits can be separated:

$$1 = \lim_{d\to\infty} \left(\frac{|\hat{S}_{n,d}|}{\frac{d^n}{n!}} + \frac{|\hat{S}_{n,d-1}|}{\frac{d^n}{n!}} \right) = \lim_{d\to\infty} \frac{|\hat{S}_{n,d}|}{\frac{d^n}{n!}} + \lim_{d\to\infty} \frac{|\hat{S}_{n,d-1}|}{\frac{d^n}{n!}}$$

Using the limit property $\lim\limits_{k\to\infty} f(k) \lim\limits_{k\to\infty} g(k) = \lim\limits_{k\to\infty} (f(k)g(k))$ if these limits exist, and using $1 = \lim\limits_{d\to\infty} \frac{d^n}{(d-1)^n}$. By multiplying this limit with $\lim\limits_{d\to\infty} \frac{|\hat{S}_{n,d-1}|}{\frac{d^n}{n!}}$ applying index transformation from d to $d+1$ it follows that:

$$1 = \lim_{d\to\infty} \left(\frac{|\hat{S}_{n,d}|}{\frac{d^n}{n!}} + \frac{|\hat{S}_{n,d-1}|}{\frac{d^n}{n!}} \right) = \lim_{d\to\infty} \frac{|\hat{S}_{n,d}|}{\frac{d^n}{n!}} + \lim_{d\to\infty} \frac{|\hat{S}_{n,d-1}|}{\frac{(d-1)^n}{n!}} = 2 \lim_{d\to\infty} \frac{|\hat{S}_{n,d}|}{\frac{d^n}{n!}}$$

So $\lim\limits_{d\to\infty} \frac{|\hat{S}_{n,d}|}{\frac{d^n}{n!}} = \frac{1}{2}$. Therefore, the leading term of $\varphi(n,d) = |\hat{S}_{n,d}|$ is $\frac{d^n}{2n!}$.

So as d becomes large, one finds $\frac{d^n}{n!} \approx |\hat{U}_{n,d}| \approx 2|\hat{S}_{n,d}| = 2\varphi(n,d)$. So the leading term of $\varphi(n,d)$ as d becomes large is $\frac{d^n}{2n!}$.

3.5 Improved method

Section 3.4 provided a method to write any general polynomial of any total degree and any number of variables as a determinant of a matrix M that is of the following form.

$$M = \begin{pmatrix} * & -1 & & & \\ * & * & \ddots & & \\ * & * & * & \ddots & \\ * & * & * & * & -1 \\ r_1 & r_2 & \dots & \dots & r_N \end{pmatrix} = \begin{pmatrix} & G & \\ r_1 & \dots & r_N \end{pmatrix}$$

Here G is a $(N-1) \times N$ matrix.

It holds that $\det M = \sum_{i=1}^{N} r_i m_i$ where m_i is a monomial from the set S in an ordered way. Denote $M^{(i,j)}$ as the minor matrix of M omitting row i and column j, and denote $G^{(j)}$ as a square matrix by taking G and omit column j. Then one can write the determinant of M by expanding the last row into minors: $\det M = \sum_{i=1}^{N} r_i \det M^{(N,i)} = \sum_{i=1}^{N} r_i (-1)^{N+i} \det G^{(i)}$.

Now choose a different set of variables $\widetilde{S}$ with $|\widetilde{S}| = \widetilde{N}$, and make the same construction. The result is a $\widetilde{N} \times \widetilde{N}$-matrix $\widetilde{M}$.

Let $P = (p_{i,j})$ with $p_{i,j} = 1$ if $i + j = \widetilde{N}$ and $p_{i,j} = 0$ otherwise. This is a permutation matrix that reverses the order of the rows and columns when multiplying. Also note that $\det P = \pm 1$, so that $\det P\widetilde{M}P = (\det P)^2 \det \widetilde{M} = \det \widetilde{M}$. Now consider the matrix $P\widetilde{M}P$ and define the right block as $\widetilde{G}$.

$$P\widetilde{M}P = \begin{pmatrix} \widetilde{r}_{\widetilde{N}} & -1 & & & \\ \vdots & * & \ddots & & \\ \vdots & * & * & \ddots & \\ \widetilde{r}_2 & * & * & * & -1 \\ \widetilde{r}_1 & * & * & * & * \end{pmatrix} = \begin{pmatrix} \widetilde{r}_{\widetilde{N}} & & \\ \vdots & & \widetilde{G} \\ \widetilde{r}_1 & & \end{pmatrix}$$

Now introduce a new matrix W of dimension $(N + \widetilde{N} - 1) \times (N + \widetilde{N} - 1)$ that is defined as follows:

$$W = \begin{pmatrix} G & O \\ K & \widetilde{G} \end{pmatrix} = \left(\begin{array}{cccc|ccc} * & -1 & & & & & \\ * & * & \ddots & & & & \\ * & * & * & -1 & & & \\ \hline k_{1,\widetilde{N}} & k_{2,\widetilde{N}} & \dots & k_{N,\widetilde{N}} & -1 & & \\ k_{1,\widetilde{N}-1} & k_{2,\widetilde{N}-1} & \ddots & \vdots & * & \ddots & \\ \vdots & \ddots & \ddots & \vdots & * & * & -1 \\ k_{1,1} & \dots & \dots & k_{N,1} & * & * & * \end{array}\right)$$

In this new definition, K is a $\widetilde{N} \times N$-matrix containing entries like the entries r_i before. If the matrices G and $\widetilde{G}$ are constructed with the sets $S = \{1 = m_1, m_2, \dots, m_N\}$ and $\widetilde{S} = \{1 = \widetilde{m}_1, \widetilde{m}_2, \dots, \widetilde{m}_{\widetilde{N}}\}$, then the claim is that $\det W = \sum_{i=1}^{N} \sum_{j=1}^{\widetilde{N}} k_{i,j} m_i \widetilde{m}_j$.

3.5.1 Proof of correctness of the construction

This section will prove the correctness of the construction. It will be proven that $\det W = \sum_{i=1}^{N} \sum_{j=1}^{\widetilde{N}} k_{i,j} m_i \widetilde{m}_j$.

Proposition 3.5.1 $\det W = \sum_{i=1}^{N} \sum_{j=1}^{\widetilde{N}} k_{i,j} m_i \widetilde{m}_j$.

Proof The proof consists of three parts. First will be proven that in $\det W$ there cannot be terms consisting of two factors from the matrix K. Next will be proven that in $\det W$ there cannot be terms without entries from matrix K. Using these results, in the last part of the proof will be proven that the remaining factors in the terms with factor $k_{i,j}$ are indeed $m_i \widetilde{m}_j$, and that the sign is always positive.

The proof will use the following simplified representation of matrix W:

$$W = \begin{pmatrix} G & O \\ K & \widetilde{G} \end{pmatrix}$$

Suppose there is a nonzero term $k_{\alpha,\beta} k_{\gamma,\delta} g$ in $\det W$ where $k_{\alpha,\beta}$ and $k_{\gamma,\delta}$ are entries from the matrix K, and where g is a product of nonzero entries from the whole matrix. Then there is a minor matrix of matrix W where the rows and columns containing $k_{\alpha,\beta}$ and $k_{\gamma,\delta}$ are omitted. Denote this matrix as $W^{\square}$. Also denote $K^{\square}$ as the matrix with the rows and columns omitted, and $G^{\square}$ as the matrix with the columns omitted, and $\widetilde{G}^{\square}$ as the matrix with the rows omitted.

By assumption, because $k_{\alpha,\beta} k_{\gamma,\delta} g$ is a nonzero term of $\det W$, $0 \neq g = \det W^{\square}$.

Matrix $\widetilde{G}$ has $\widetilde{N}$ rows and $\widetilde{N} - 1$ columns. In the minor expansion, two rows of this matrix are omitted. This means that there are $\widetilde{N} - 2$ rows and $\widetilde{N} - 1$ columns in $\widetilde{G}^{\square}$. Because the dimension of the row space of a matrix is the same as the dimension of the column-space of a matrix, the number of rows is an upper bound for the dimension of the column space. This implies that the column space of $\widetilde{G}^{\square}$ is at most $(\widetilde{N} - 2)$-dimensional. Also the column space of $\begin{pmatrix} O \\ \widetilde{G}^{\square} \end{pmatrix}$ is $(\widetilde{N} - 2)$-dimensional, because the zero-matrix cannot increase the dimension. This implies that the columns of the matrix $\begin{pmatrix} O \\ \widetilde{G}^{\square} \end{pmatrix}$ are dependent, so the columns of the bigger matrix $W^{\square} = \begin{pmatrix} G^{\square} & O \\ K^{\square} & \widetilde{G}^{\square} \end{pmatrix}$ are dependent. Dependency of the columns implies that $g = \det W^{\square} = 0$, which contradicts the assumption that $g \neq 0$.

For this reason, it's not possible that there are nonzero terms $k_{\alpha,\beta} k_{\gamma,\delta} g$ in $\det W$. **(Result 1)**

Suppose there is a term in $\det W$ without any entry from the matrix K. In that case, the entry occurs in the matrix $W' = \begin{pmatrix} G & O \\ O & \widetilde{G} \end{pmatrix}$ as well. Denote $\widetilde{G}^{(i)}$ as the matrix $\widetilde{G}$ with the rows and columns 1 row i omitted. Note that $\widetilde{G}^{(i)}$ is of the same form as $\widetilde{G}$: It has diagonal entries

-1, and no entries above the diagonal.

$$\widetilde{G} = \overbrace{\begin{pmatrix} -1 & & \\ * & \ddots & \\ * & * & -1 \\ * & * & * \end{pmatrix}}^{\widetilde{N}-1 \text{ columns}} \quad , \quad \widetilde{G}^{(i)} = \overbrace{\begin{pmatrix} -1 & & \\ * & \ddots & \\ * & * & -1 \\ * & * & * \end{pmatrix}}^{\widetilde{N}-1-i \text{ columns}}$$

Now iteratively apply minor expansion to the top-left entry of the bottom-right block. The value of this entry will always be -1, because it's a diagonal entry of $\widetilde{G}$. In matrix W this entry is at position $(N, N+1)$, so the permutation sign will be -1 at all times. Now iteratively apply expansion into minors:

$$\begin{aligned} \det W' = \det \begin{pmatrix} G & O \\ O & \widetilde{G} \end{pmatrix} &= -1 \cdot (-1)^{N+N+1} \cdot \det \begin{pmatrix} G & O \\ O & \widetilde{G}^{(1)} \end{pmatrix} = \det \begin{pmatrix} G & O \\ O & \widetilde{G}^{(1)} \end{pmatrix} \\ &= -1 \cdot (-1)^{N+N+1} \cdot \det \begin{pmatrix} G & O \\ O & \widetilde{G}^{(2)} \end{pmatrix} = \det \begin{pmatrix} G & O \\ O & \widetilde{G}^{(2)} \end{pmatrix} \\ &= \cdots = \det \begin{pmatrix} G & O \\ O & \widetilde{G}^{(\tilde{N}-2)} \end{pmatrix} = \det \left(\begin{array}{c|c} G & \underline{0} \\ \hline \underline{0}^T & -1 \\ \underline{0}^T & * \end{array} \right) \\ &= -1 \cdot (-1)^{N+N+1} \cdot \det \begin{pmatrix} G \\ \underline{0}^T \end{pmatrix} = \det \begin{pmatrix} G \\ \underline{0}^T \end{pmatrix} = 0 \end{aligned}$$

In the last step it's found that $\det W' = \det \begin{pmatrix} G \\ \underline{0}^T \end{pmatrix}$. Because there is a zero row in the square matrix $\begin{pmatrix} G \\ \underline{0}^T \end{pmatrix}$, the determinant is 0. So if $K = O$, then $\det W = 0$. Because $\det W' = 0$ and every term of $\det W$ without a factor from the entries of K is a term from $\det W'$, it follows that there is no term from $\det W$ with no factor from the matrix K. **(Result 2)**

Combining the two results yields that every term from $\det W$ contains exactly one entry from the matrix K. What follows is an expression for $\det W$:

$$\det W = \sum_{i=1}^{N} \sum_{j=1}^{\widetilde{N}} k_{i,j} \sigma(i,j) \det W^{(i,j)} \tag{3.2}$$

In this expression, $W^{(i,j)}$ is the matrix W with the row and column containing the entry $k_{i,j}$ omitted. The entry $k_{i,j}$ is located at row $N + \widetilde{N} - j$ and at column i. Therefore, the permutation sign is $\sigma(i,j) = (-1)^{N+\widetilde{N}-j+i} = (-1)^{N+\widetilde{N}+i+j}$. Note that $\det W^{(i,j)}$ does not contain a factor that is an entry of the matrix K because of result 1: There is no term in $\det W$ that contains two entries of the matrix K as factor.

Therefore $\det W^{(i,j)} = \det \begin{pmatrix} G^* & O \\ K^* & \widetilde{G}^* \end{pmatrix} = \det \begin{pmatrix} G^* & O \\ O & \widetilde{G}^* \end{pmatrix} = \det((W')^{(i,j)})$. In this notation, G^* is matrix G with the column of $k_{i,j}$ omitted and $\widetilde{G}^*$ is matrix $\widetilde{G}$ with the row of $k_{i,j}$ omitted. Because G is a $((N-1) \times N)$-matrix and $\widetilde{G}$ a $(\widetilde{N} \times (\widetilde{N}-1))$-matrix, the matrices G^* and $\widetilde{G}^*$

are both square matrices. That means that the matrix $(W')^{(i,j)}$ is in block-diagonal form, so $\det(W')^{(i,j)} = \det G^* \det \widetilde{G}^*$.

Note that $\det G^* = (-1)^{N+i} \det \begin{pmatrix} G \\ e_i^T \end{pmatrix}$, and that $\det \widetilde{G}^* = (-1)^{(\widetilde{N}-j+1)+1} \det \begin{pmatrix} e_{\widetilde{N}-j+1} & \widetilde{G} \end{pmatrix}$. These ways of writing the matrices is in particular useful, because these are variants of the construction discussed in Section 3.4. The result is that $\det G^* = (-1)^{N+i} \det \begin{pmatrix} G \\ e_i^T \end{pmatrix} = (-1)^{N+i} m_i$, because the matrix is in the form of the generalized method of Plestenjak and Hochstenbach, so the determinant is the inner product of the ordered monomial list as vector with the ith basic vector, which gives m_i.

Transposing $\begin{pmatrix} e_{\widetilde{N}-j+1} & \widetilde{G} \end{pmatrix}$ and applying the reverse permutation on the rows and columns does not change the determinant, and also yields the form of the construction by Plestenjak and Hochstenbach. Using these observations the determinant of $\widetilde{G}^*$ can be derived as follows:

$$\begin{aligned} \det \widetilde{G}^* &= (-1)^{(\widetilde{N}-j+1)+1} \det \begin{pmatrix} e_{\widetilde{N}-j+1} & \widetilde{G} \end{pmatrix} = (-1)^{\widetilde{N}+j} \det \left(P \begin{pmatrix} e_{\widetilde{N}-j+1} & \widetilde{G} \end{pmatrix}^T P \right) \\ &= (-1)^{\widetilde{N}+j} \widetilde{m}_j = (-1)^{\widetilde{N}+j} \widetilde{m}_j \end{aligned}$$

In this notation, P is the reverse $\widetilde{N} \times \widetilde{N}$ permutation matrix $P = (p_{i,j})$ with $p_{i,j} = 1$ if $i + j = \widetilde{N}$ and 0 otherwise.

What follows is that $\det G^* = (-1)^{N+i} m_i$ and $\det \widetilde{G}^* = (-1)^{\widetilde{N}+j} \widetilde{m}_j$. Filling this in in the block expression yields:

$$\det W^{(i,j)} = \det(W')^{(i,j)} = \det G^* \det \widetilde{G}^* = (-1)^{N+\widetilde{N}+i+j} m_i \cdot \widetilde{m}_j = \sigma(i,j) m_i \cdot \widetilde{m}_j$$

Using that $\sigma(i,j) = \pm 1$, it follows that $(\sigma(i,j))^2 = 1$. Substituting this in Equation 3.2 yields:

$$\det W = \sum_{i=1}^{N} \sum_{j=1}^{\widetilde{N}} k_{i,j} \sigma(i,j) \det W^{(i,j)} = \sum_{i=1}^{N} \sum_{j=1}^{\widetilde{N}} \sigma(i,j) k_{i,j} \cdot \sigma(i,j) m_i \widetilde{m}_j = \sum_{i=1}^{N} \sum_{j=1}^{\widetilde{N}} k_{i,j} m_i \widetilde{m}_j$$

This completes the proof. □

3.5.2 Asymptotic Analysis

The construction of the improved method relies on the generalized construction that is discussed in Section 3.4. If a polynomial p of degree d in n variables is given, a possibility is to split up the variables in two groups. Let the monomials $x_1, \ldots, x_k$ be part of the first group, and the monomials $x_{k+1}, \ldots, x_n$ be part of the second group. This section will discuss some possible directions and the related asymptotic behavior. However, no statements about minimality under this construction will be made.

A possible construction follows. Choose set $S = \{x_1^{i_1} \ldots x_k^{i_k} \mid i_1 + \cdots + i_k \leq d\}$ and $\widetilde{S} = \{x_{k+1}^{i_{k+1}} \ldots x_n^{i_n} \mid i_{k+1} + \cdots + i_n \leq d\}$. Then $N = |S| = \binom{k+d}{d}$ and $\widetilde{N} = |\widetilde{S}| = \binom{n-k+d}{d}$. Then the construction of the improved method gives a $(N + \widetilde{N} - 1) \times (N + \widetilde{N} - 1)$-matrix W that has the polynomial p as determinant. So the square matrix will have the size $\binom{k+d}{d} + \binom{n-k+d}{d} - 1$.

It is mainly interesting how well this representation performs for a higher number of variables and for a higher degree. One can write the binomial terms as products as follows:

$$\binom{k+d}{d} = \frac{(k+d)(k-1+d)\dots(1+d)}{k!} = \frac{\prod_{l=1}^{k}(d+l)}{k!}$$

And likewise $\binom{n-k+d}{d} = \frac{1}{(n-k)!}\prod_{l=1}^{n-k}(d+l)$. From this it follows that the size of the matrix is $\frac{\prod_{l=1}^{k}(d+l)}{k!} + \frac{\prod_{l=1}^{n-k}(d+l)}{(n-k)!} - 1$, so the order with respect to d is determined by the faster growing term: $\frac{d^k}{k!}$ or $\frac{d^{n-k}}{(n-k)!}$. The optimal situation occurs when the exponent of the order-term is minimal. This occurs when $k \approx \frac{1}{2}n$.

In the case where the number of variables n is even, choosing $k = \frac{1}{2}n$ is simple: One can pick one half of the monomials for the set S, and the other half for the set $\widetilde{S}$. In the case where n is odd, a possibility is choosing $\lfloor \frac{1}{2}n \rfloor$ monomials for the first group and $\lceil \frac{1}{2}n \rceil$ monomials for the second group.

Because every monomial that occurs in the polynomial is in the set $S \cdot \widetilde{S}$ (by construction), the matrix K can be chosen as a matrix that only contains the coefficients. This provides a simple construction, because there are no terms in K that are linear in the variables. The downside is that this yields matrices of bigger size. It is possible to reduce the size of the matrix, requiring terms in the matrix K that are affine-linear in the variables. In that case the sets S and $\widetilde{S}$ should be chosen such that every monomial of p can be created by multiplying a variable $x_1, \dots, x_n$ by some element from $S \cdot \widetilde{S}$.

A simple improvement will be by leaving out the highest-order terms. Then S and $\widetilde{S}$ can be chosen as follows:

$$S = \{x_1^{i_1} \dots x_k^{i_k} \mid i_1 + \dots + i_k \le d-1\} \qquad \widetilde{S} = \{x_{k+1}^{i_{k+1}} \dots x_n^{i_n} \mid i_{k+1} + \dots + i_n \le d-1\}$$

These sets have the cardinalities $\binom{k+d-1}{d-1}$ and $\binom{n-k+d-1}{d-1}$ respectively. Since $\binom{k+d-1}{d-1} = \frac{\prod_{l=1}^{k} d-1+l}{k!}$ this will offer a matrix of smaller size, but it will have the same asymptotic order term $\frac{d^k}{k!}$.

Another possibility may be choosing either the set S or the set $\widetilde{S}$ as in Section 3.4. It is not possible to choose both sets as described in this construction, because there may be a monomial of which the $x_1, \dots, x_k$-part may not be in S and the $x_{k+1}, \dots, x_n$-part may not be in $\widetilde{S}$. As discussed in Section 3.4.2, this yields that the cardinality of the augmented set will be of order $\frac{d^k}{2k!}$ in the case considering set S. Even though the factor $\frac{1}{2}$ is a considerable improvement over the first proposal, the asymptotic order will remain the same because the exponent of d will be unchanged.

Compared to the generalized method of Section 3.4, the improved method provides a considerable improvement asymptotically to d. This improved method will give a matrix representation where the matrix is of asymptotic size $\mathcal{O}(d^{\lceil \frac{n}{2} \rceil})$, whereas the generalized method provided a matrix of asymptotic size $\mathcal{O}(d^n)$.

Chapter 4

Lower bounds for the matrix dimensions

This chapter focuses on finding lower bounds for matrix dimensions for polynomials of a certain degree and with a certain number of variables. In the first section polynomials of a certain degree and certain number of variables are considered, and a tight lower bound for these polynomials will be derived. In the second part an upper bound is derived.

4.1 Global lower bound

Let p be a polynomial with n variables $x_1, \dots, x_n$ and of degree d: $p \in \mathbb{C}[x_1, \dots, x_n]_{\leq d}$. Denote $p = \sum_{\alpha \in \mathbb{N}_0^n, |\alpha| \leq d} c_\alpha x^\alpha$, where c_α are the coefficients of the polynomial and $x = (x_1, \dots, x_n)$.

Let M be a $N \times N$ matrix with entries that are affine linear in the variables $x_1, \dots, x_n$ and in the coefficients c_α. Writing $\mathcal{V} = \{x_1, \dots, x_n\}$ as the set of variables and $\mathcal{D} = \{c_\alpha \mid \alpha \in \mathbb{N}_0, |\alpha| \leq d\}$ as the set of coefficients, then $M \in \mathbb{C}[\mathcal{D}]_{\leq 1}[\mathcal{V}]_{\leq 1}^{N \times N}$. The property of M is that $\det M = p$.

Consider the case where the polynomial p is chosen, such that the coefficients of p are fixed. Then the variables of the matrix M are the variables of the polynomial p only: $x_1, \dots, x_n$. So $M \in \mathbb{C}[x_1, \dots, x_n]_{\leq 1}^{N \times N}$.

In this section a lower bound for the size of matrix M is found. Since M is a $N \times N$ square matrix, the size of the matrix will be referred to as N.

In the determinant of a $N \times N$-matrix, each term is a product of N entries. Each entry in the matrix M is a polynomial of maximal degree one in $x_1, \dots, x_n$. Since p is a polynomial of degree d, there must be terms that are products of d entries of M. So a lower bound for the size N of the $N \times N$-matrix M is $N \geq d$. This is a simple initial lower bound.

Since p is a polynomial of degree d and with n variables, there are $\binom{n+d}{d}$ terms in this polynomial. Each term has one coefficient, so p has $\binom{n+d}{d}$ degrees of freedom in the choice of the coefficients. $p = \det M$, so $\det M$ should have at least $\binom{n+d}{d}$ degrees of freedom as well.

Therefore M has at least $\binom{n+d}{d}$ degrees of freedom. Summarizing:

$$\begin{aligned}&\text{Number of degrees of freedom in } M\\ &\geq \text{Number of degrees of freedom in } \det M\\ &= \text{Number of degrees of freedom in } p\\ &= \text{Number of terms in } p = \binom{n+d}{n}\end{aligned}$$

Since M is a $N \times N$-matrix, it has N^2 matrix entries. Each entry is of the form $\beta_0 + \sum_{i=1}^{n} \beta_i x_i$, where $\beta_i \in \mathbb{C}$ since p is fixed. Therefore, each entry has $n+1$ free choices, namely one choice per monomial $1, x_1, \ldots, x_n$, corresponding to $\beta_0, \ldots, \beta_n$. What follows is that the number of degrees of freedom in M is $(n+1)N^2$.

For this reason, $(n+1)N^2 \geq \binom{n+d}{d}$ gives a lower bound for the size of the matrix M. Solving this expression to N yields the following lower bound for N:

$$N \geq \sqrt{\frac{\binom{n+d}{d}}{n+1}}$$

4.1.1 Improvement of the lower bound

Consider p as a polynomial of n variables and of degree d. Writing $\mathcal{V} = \{x_1, \ldots, x_n\}$ as the variables of p, then $p \in \mathbb{C}[\mathcal{V}]_{\leq d}$. Let M be a $N \times N$-matrix that is affine linear in the variables and affine linear in the coefficients of p, such that $\det M = p$.

Define $\psi : \mathbb{C}[\mathcal{V}]_{\leq 1}^{N \times N} \to \mathbb{C}[\mathcal{V}]$ as $\psi(M) = \det M$. It is required that for each polynomial $p \in \mathbb{C}[\mathcal{V}]_{\leq d}$ there exists a $N \times N$-matrix M such that $\det M = p$. Since $p \in \mathbb{C}[\mathcal{V}]_{\leq d}$, and $\det M = p$ is required, it must hold that $\mathbb{C}[\mathcal{V}]_{\leq d} \subseteq \operatorname{im} \psi$.

It holds that $\dim \mathbb{C}[\mathcal{V}]_{\leq d} = \binom{n+d}{d}$, because that is the number of monomials of p, and each monomial has a coefficient that can be picked independent of the other coefficients. From this follows the following lower bound for $\dim \operatorname{im} \psi$:

$$\binom{n+d}{d} = \dim \mathbb{C}[\mathcal{V}]_{\leq d} \leq \dim \operatorname{im} \psi$$

For all $S \in \mathcal{SL}_N(\mathbb{C})$ it holds that $\det MS = \det M$, so $\psi(M) = \psi(MS)$. It holds that $\dim \mathcal{SL}_N(\mathbb{C}) = N^2 - 1$. Using this result, one can find a lower bound for the dimension of the set $\psi^{-1}(\psi(M)) = \{K \in \mathbb{C}[x]_{\leq 1}^{N \times N} \mid \det K = \det M\}$ as follows:

$$\dim \psi^{-1}(\psi(M)) \geq \dim \mathcal{SL}_N(\mathbb{C}) = N^2 - 1$$

From this it follows that there exists a matrix space in which every matrix has the same determinant that has a dimension that is at least $N^2 - 1$. Therefore, $\dim \operatorname{im} \psi \leq \dim \mathbb{C}[x]_{\leq 1}^{N \times N} - (N^2 - 1)$.

It holds that $\dim \mathbb{C}[x]_{\leq 1}^{N \times N} = (n+1)N^2$, because every entry of the matrix has dimension $n+1$, and because there are N^2 entries.

Combining all results yields the following estimation:

$$\binom{n+d}{d} = \dim \mathbb{C}\left[\mathcal{V}\right]_{\leq d} \leq \dim \operatorname{im} \psi \leq \dim \mathbb{C}\left[x\right]_{\leq 1}^{N \times N} - (N^2-1) = (n+1)N^2 - (N^2-1) = nN^2+1$$

In short, the estimate is as follows: $\binom{n+d}{d} \leq nN^2 + 1$. Solving this expression to N yields the following lower bound for N:

$$N \geq \sqrt{\frac{\binom{n+d}{d} - 1}{n}}$$

4.2 Tight lower bounds for certain cases

The previous section discussed lower bounds for matrices of which the determinants are general polynomials of n variables and of degree d. This section will provide tight upper bounds for matrices of minimal size for polynomials of certain degrees and with a certain number of variables. This will be done for degree-one polynomials with any number of variables, degree-two bivariate polynomials and for degree-two trivariate polynomials.

4.2.1 Degree-one multivariate polynomials

Let $p(x_1, \ldots, x_n) = c_0 + \sum_{k=1}^{n} c_k x_k$ be a polynomial of maximal degree one and with n variables. In this thesis, constructions of matrices are considered where the entries of the matrix are affine linear in the coefficients and affine linear in the variables.

In the case of polynomials of degree one, the polynomial itself is already an expression that is affine linear in the variables and affine linear in the coefficients. For this reason, the polynomial $p(x_1, \ldots, x_n)$ satisfies the requirements for being a matrix entry.

A straightforward construction follows by choosing $M = (p(x_1, \ldots, x_n))$, which is an 1×1-matrix. For an 1×1-matrix it holds that the determinant is same as the entry itself. Therefore, $\det M = p(x_1, \ldots, x_n)$. It is obvious that this is the minimal case, as an 1×1-matrix is a matrix of minimal size.

4.2.2 Degree-two bivariate polynomials

In Section 3.3 the construction by Plestenjak and Hochstenbach is discussed. For degree-two bivariate polynomials, picking the monomial set $\{1, x, y\}$ is the minimal set such that all monomials up to degree two can be created by multiplying monomials in this set with 1, x or y. The construction of Plestenjak and Hochstenbach provides a matrix that has the general degree-two bivariate polynomial as determinant, which has the dimensions equal to the cardinality of that set. So their construction provides an upper bound of a matrix that is 3×3.

It is interesting to check if it is possible to create a 2×2-matrix of which the entries are affine linear in the coefficients and affine linear in the variables of the polynomial, such that the determinant of this matrix is also the general degree-two bivariate polynomial.

Let $p(x,y) = \sum_{i=0}^{2}\sum_{j=0}^{2-i} d_{i,j}x^iy^j$ and suppose there exists a matrix $M(x,y) = M_0(x,y) + M_1(x,y)$ such that $\det M(x,y) = p(x,y)$, where M_0 is independent of the coefficients of p, and where $M_1(x,y)$ only contains terms dependent on the coefficients of p.

Using Proposition 3.1.1 it follows that $\det M_0 = 0$ for all values of x and y. Using proposition 3.1.2 it follows that $\det M_1(x,y) = 0$ for all values of x and y. Using proposition 3.1.3, it follows that $\operatorname{rank} M_0 = 1$ for all values of x and y.

$\det M_0(x,y) = 0$ and $\operatorname{rank} M_0(x,y) = 1$ for all x and y, so especially $\det M_0(0,0) = 0$ and $\operatorname{rank} M_0(0,0) = 1$. So there exists matrices $S,T \in \mathcal{SL}_2(\mathbb{C})$ such that $SM_0(0,0)T = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. Then the matrix $SM_0(x,y)T = \begin{pmatrix} 1+a_1x+b_1y & a_2x+b_2y \\ a_3x+b_3y & a_4x+b_4y \end{pmatrix}$. Since $S,T \in \mathcal{SL}_2(\mathbb{C})$, it holds that $\det S = \det T = 1$, so $\det SM(x,y)T = \det S \det M(x,y) \det T = \det M(x,y)$. Therefore one can assume without loss of generality that $M_0(x,y)$ itself is of the form presented above.

From now on, let $M_0(0,0) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$.

The determinant of this matrix must be zero, so:

$$\det M_0(x,y) = (1+a_1x+b_1y)(a_4x+b_4y) - (a_2x+b_2y)(a_3x+b_3y) = 0$$

From this it follows that the linear part $a_4x+b_4y = 0$ and hence $(a_2x+b_2y)(a_3x+b_3y) = 0$. It does not matter if one chooses $a_2x+b_2y = 0$ or $a_3x+b_3y = 0$ because of symmetry: transposing the matrix gives the other case. Without loss of generality, suppose $a_3x+b_3y = 0$ and $a_4x+b_4y = 0$. With these reductions of the problem applied, M_0 follows:

$$M_0(x,y) = \begin{pmatrix} 1+a_1x+b_1y & a_2x+b_2y \\ 0 & 0 \end{pmatrix}$$

Define $M_1(x,y) = \begin{pmatrix} m_1 & m_2 \\ m_3 & m_4 \end{pmatrix}$, where m_1, m_2, m_3 and m_4 are linear in the coefficients and in 1, x and y. Then $M(x,y) = \begin{pmatrix} 1+m_1+a_1x+b_1y & m_2+a_2x+b_2y \\ m_3 & m_4 \end{pmatrix}$. It must hold that $\det M_1(x,y) = 0$, so $m_1m_4 - m_2m_3 = 0$. Now compute $\det M(x,y)$:

$$\begin{aligned}\det M(x,y) &= m_4(1+m_1+a_1x+b_1y) - m_3(m_2+a_2x+b_2y) \\ &= m_4(1+a_1x+b_1y) - m_3(a_2x+b_2y) + m_1m_4 - m_2m_3 \\ &= m_4(1+a_1x+b_1y) - m_3(a_2x+b_2y)\end{aligned}$$

Define $\Omega : \mathbb{C}[x,y]_{\leq 1}^2 \to \mathbb{C}[x,y]$ as $\Omega(m_3,m_4) = \det M$. The dimension of the domain of Ω is $\dim \mathbb{C}[x,y]_{\leq 1}^2 = 2\dim \mathbb{C}[x,y]_{\leq 1} = 2\cdot 3 = 6$.

Note that choosing $m_3 = 1+a_1x+b_1y$ and $m_4 = a_2x+b_2y$ yields $\Omega(1+a_1x+b_1y, a_2x+b_2y) = 0$, so $(1+a_1x+b_1y, a_2x+b_2y) \in \ker(\Omega)$. For this reason, $\dim\ker(\Omega) \geq 1$.

Using rank nullity it must hold that that $\dim \operatorname{im} \Omega = \dim \mathbb{C}[x,y]_{\leq 1}^2 - \dim\ker\Omega$. It holds that $\dim\ker\Omega \geq 1$ and $\dim \mathbb{C}[x,y]_{\leq 1}^2 = 6$. From this it follows that $\dim \operatorname{im} \Omega = [x,y]_{\leq 1}^2 - \dim\ker\Omega \leq 6-1 = 5$. This provides an upper bound for the image of Ω, namely $\dim \operatorname{im} \Omega \leq 5$.

The polynomial $p(x,y)$ has 6 coefficients, so using that $\det M = p$ provides a lower bound for the image of Ω. This lower bound is $\dim \operatorname{im} \Omega \geq 6$.

This shows a contradiction, namely $6 \leq \dim \operatorname{im} \Omega \leq 5$. Therefore it is not possible to write a general degree-two bivariate polynomial p as a determinant of a 2×2 matrix that is affine linear in the variables and affine linear in the coefficients of the polynomial.

This also shows that the construction by Plestenjak and Hochstenbach is optimal for the degree-two bivariate case.

4.2.3 Degree-two trivariate polynomials

For any degree-two trivariate polynomial, the construction of Plestenjak and Hochstenbach provides a 4×4-matrix M of which the determinant is the considered polynomial. This section will show that a 4×4-matrix for any degree-two trivariate polynomial is minimal. As in the previous section, this will be done by showing that it is not possible to do it with a matrix that has a smaller size.

Consider the polynomial p that is of degree two in three variables x, y and z.

$$p(x,y,z) = \sum_{i_1+i_2+i_3 \leq 2,\ i_1,i_2,i_3 \in \mathbb{N}_0} c_{i_1,i_2,i_3} x^{i_1} y^{i_2} z^{i_3}$$

Define $\mathcal{V} = \{x, y, z\}$ as the set of the variables of the polynomial p, and $\mathcal{D} = \{d_{i,j,k} | i, j, k \in \mathbb{N}_0, i + j + k \leq 2\}$ as the set of coefficients.

Let $M \in \mathbb{C}\left[\mathcal{D}\right]_{\leq 1} \left[\mathcal{V}\right]_{\leq 1}^{3\times 3}$ be a 3×3-matrix such that $\det M = p$. Like before, let $M_0 \in \mathbb{C}\left[\mathcal{V}\right]_{\leq 1}^{3\times 3}$ and $M_1 \in \mathbb{C}\left[\mathcal{D}\right]_{=1} \left[\mathcal{V}\right]_{\leq 1}^{3\times 3}$, such that $M = M_0 + M_1$.

Using proposition 3.1.1 one finds that $\det M_0(x, y, z) = 0$ for all values for x, y and z. Since M_0 is a 3×3-matrix there exist $S, T \in \mathcal{SL}_3(\mathbb{C})$ such that $SM_0(x, y, z)T$ is of one of the following types:

$$(1) \begin{pmatrix} * & * & 0 \\ * & * & 0 \\ * & * & 0 \end{pmatrix} \text{or} \begin{pmatrix} * & * & * \\ * & * & * \\ 0 & 0 & 0 \end{pmatrix} \qquad (2) \begin{pmatrix} * & * & * \\ 0 & 0 & * \\ 0 & 0 & * \end{pmatrix} \qquad (3) \begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{pmatrix}$$

These types are listed at page 265 of the report by P. Fillmore et. al [6]. The first two types are examples of 3×3 compression spaces. The third type is a skew-symmetric matrix. If $M = -M^T$ one finds for matrices of size $N \times N$ that $\det M = (-1)^N \det M$, which implies that $\det M = 0$ if N is odd. Because matrix spaces of size 3×3 are considered, the skew-symmetric case is a separate case in the space of matrices with determinant zero.

So suppose x, y and z are fixed, then one can assume without loss of generality that $M_0(x, y, z)$ is either of type (1), (2) or (3). In the next sections a contradiction for all cases will be derived.

Analysis where M_0 is of the type (1)

Suppose the matrix M_0 is of type (1) as discussed above. Because transposing the matrix in computing the determinant of M, $\det M = \det M^T$, it suffices to show non-existence for one of the two variants. Choose the variant where M_0 can be transformed into a form containing one zero-column.

Write M_0 as follows:

$$M_0 = \begin{pmatrix} n_{1,1} & n_{1,2} & 0 \\ n_{2,1} & n_{2,2} & 0 \\ n_{3,1} & n_{3,2} & 0 \end{pmatrix}$$

In this notation, the entries $n_{i,j} \in \mathbb{C}\left[\mathcal{V}\right]_{\leq 1}$.

Next the form of M_1 will be discussed when M_0 is as above. Let M_1 be defined as follows:

$$M_1 = \begin{pmatrix} m_{1,1} & m_{1,2} & m_{1,3} \\ m_{2,1} & m_{2,2} & m_{2,3} \\ m_{3,1} & m_{3,2} & m_{3,3} \end{pmatrix}$$

In this notation, the entries $m_{i,j} \in \mathbb{C}\left[\mathcal{D}\right]_{=1}\left[\mathcal{V}\right]_{\leq 1}$.

Because the last column of M_0 is the zero-column, the last column of $M := M_0 + M_1$ is exactly the last column of M_1. Denote $M^{(i,j)}$ as the minor matrix of M that is found by omitting row i and column j.

Compute $\det M$ by expanding the determinant into minors for the last column.

$$p(x,y,z) := p = \det M = m_{1,3} \det M^{(1,3)} - m_{2,3} \det M^{(2,3)} + m_{3,3} \det M^{(3,3)}$$

It holds that $m_{1,3}, m_{2,3}, m_{3,3} \in \operatorname{span}(\mathcal{D} \cdot (\mathcal{V} \cup \{1\})) = \operatorname{span}(\mathcal{D} \cdot \{1, x, y, z\}) = \operatorname{span}(\mathcal{D})$, because $x, y, z \in \mathbb{C}$ are fixed.

Also, the polynomial $p(x,y,z)$ is a linear combination of the entries in $\mathcal{D}$, so $p(x,y,z) \in \operatorname{span}(\mathcal{D})$.

For all $x, y, z \in \mathbb{C}$ there exists a degree-two trivariate polynomial p such that $p(x,y,z) \neq 0$, so for the fixed $x, y, z \in \mathbb{C}$ that are chosen, there exists coefficients such that $p(x,y,z) \neq 0$.

Now split up the minors in two parts. One part is the part independent of $\mathcal{D}$ generated by $\det M_0^{(i,3)}$. The other is the part dependent on $\mathcal{D}$, generated by $\left(\det M^{(i,3)} - \det M_0^{(i,3)}\right)$. One finds:

$$\begin{aligned} 0 \neq p(x,y,z) &= \det M(x,y,z) = m_{1,3} \det M^{(1,3)} - m_{2,3} \det M^{(2,3)} + m_{3,3} \det M^{(3,3)} \\ &= \sum_{i=1}^{3} \left((-1)^{i+3} m_{i,3} \det M^{(i,3)}\right) \\ &= \sum_{i=1}^{3} \left((-1)^{i+3} m_{i,3} \det M_0^{(i,3)}\right) + \sum_{i=1}^{3} \left((-1)^{i+3} m_{i,3} \left(\det M^{(i,3)} - \det M_0^{(i,3)}\right)\right) \end{aligned}$$

The first sum is element of $\operatorname{span}(\mathcal{D})$. The second sum is element of $\operatorname{span}(\mathcal{D} \cdot \mathcal{D})$, because each term of this sum is multiplied by $m_{i,3}$ for some $i \in \{1,2,3\}$. Because $p(x,y,z)$ is a linear combination of $\mathcal{D}$, the second sum must equal zero.

Without loss of generality, this can be achieved by letting $M^{(i,3)} = M_0^{(i,3)}$. for $i \in \{1,2,3\}$, so choose $M_1^{(i,3)} = O$ for $i \in \{1,2,3\}$. From this it follows that $m_{i,j} = 0$ if $j = 1$ or $j = 2$. So one finds that without loss of generality M_1 is as follows:

$$M_1 = \begin{pmatrix} 0 & 0 & m_{1,3} \\ 0 & 0 & m_{2,3} \\ 0 & 0 & m_{3,3} \end{pmatrix}$$

Then the following form is found for M, where $n_{i,j} \in \mathbb{C}[\mathcal{V}]_{\leq 1}$ and $m_{i,j} \in \mathbb{C}[\mathcal{D}]_{=1}[\mathcal{V}]_{\leq 1}$. For simplicity, re-write $m_{i,j} = m_i$.

$$M = \begin{pmatrix} n_{1,1} & n_{1,2} & m_{1,3} \\ n_{2,1} & n_{2,2} & m_{2,3} \\ n_{3,1} & n_{3,2} & m_{3,3} \end{pmatrix} = \begin{pmatrix} n_{1,1} & n_{1,2} & m_1 \\ n_{2,1} & n_{2,2} & m_2 \\ n_{3,1} & n_{3,2} & m_3 \end{pmatrix}$$

What is left is the following expression for $p(x, y, z) = \det M(x, y, z)$:

$$p(x, y, z) = \det M(x, y, z) = m_1 \det M_0^{(1,3)} - m_2 \det M_0^{(2,3)} + m_3 \det M_0^{(3,3)}$$

Define the map $\Omega : \mathbb{C}[x, y, z]^3_{\leq 1} \to \mathbb{C}[x, y, z]$ as follows:

$$\Omega(m_1, m_2, m_3) = m_1 \det M_0^{(1,3)} - m_2 \det M_0^{(2,3)} + m_3 \det M_0^{(3,3)}$$

Note that $p = \det M = \Omega(m_1, m_2, m_3)$.

Because $\dim \mathbb{C}[x, y, z]^3_{\leq 1} = 3 \dim \mathbb{C}[x, y, z]_{\leq 1} = 3 \cdot 4 = 12$, it holds that $\dim \operatorname{im} \Omega + \dim \ker \Omega = 12$.

It holds that $\Omega(n_{1,1}, n_{2,1}, n_{3,1}) = 0$ and $\Omega(n_{1,2}, n_{2,2}, n_{3,2}) = 0$. More specifically any linear combination can be used: $\Omega(\lambda_1 n_{1,1} + \mu_1 n_{1,2}, \lambda_2 n_{2,1} + \mu_2 n_{2,2}, \lambda_3 n_{3,1} + \mu_3 n_{3,2}) = 0$ with $\lambda_i, \mu_i \in \mathbb{C}$. This can be easily verified by considering the matrix $M = \begin{pmatrix} n_{1,1} & n_{1,2} & m_1 \\ n_{2,1} & n_{2,2} & m_2 \\ n_{3,1} & n_{3,2} & m_3 \end{pmatrix}$, because picking the last column as a linear combination of the first two columns implies that the determinant is zero due to dependence. Therefore the entries $(n_{1,1}, n_{2,1}, n_{3,1})$ and $(n_{1,2}, n_{2,2}, n_{3,2})$ are elements of the kernel of Ω.

By Proposition 3.1.3 it must hold that $\operatorname{rank} M_0 = 2$ for all $x, y, z \in \mathbb{C}$, so it must hold that $(n_{1,1}, n_{2,1}, n_{3,1})$ and $(n_{1,2}, n_{2,2}, n_{3,2})$ are independent, so the linear span over $\mathbb{C}$ defined by $V = \langle (n_{1,1}, n_{2,1}, n_{3,1}), (n_{1,2}, n_{2,2}, n_{3,2}) \rangle_{\mathbb{C}}$ has dimension 2, and $V \subseteq \ker(\Omega)$. For this reason, $\dim \ker(\Omega) \geq \dim V = 2$.

Combining this result with $\dim \operatorname{im} \Omega + \dim \ker \Omega = 12$ yields $\dim \operatorname{im} \Omega \leq 10$.

Assuming for every trivariate polynomial of total degree 2 p there exists a matrix M with entries linear in the coefficients and in the variables of p. Then the image of Ω should contain $\mathbb{C}[x, y, z]_{\leq 2}$:

$$\operatorname{im} \Omega \supseteq \mathbb{C}[x, y, z]_{\leq 2}$$

It holds that $\dim \mathbb{C}[x, y, z]_{\leq 2} = 10$. This holds, because there are 10 monomials in every trivariate polynomial of degree two, and every monomial has a coefficient: $|\{x^i y^j z^k : i + j + k \leq 2, i, j, k \in \mathbb{N}_0\}| = \binom{4+2-1}{2} = 10$. Therefore, $\dim \mathbb{C}[x, y, z]_{\leq 2} = 10$.

It is found that $\dim \operatorname{im} \Omega \leq 10$, and also that $\operatorname{im} \Omega \supseteq \mathbb{C}[x, y, z]_{\leq 2}$. This can only hold if equality of the dimension of Ω holds: $\dim \Omega = 10$. In that case it follows that $\operatorname{im} \Omega = \mathbb{C}[x, y, z]_{\leq 2}$.

Now suppose that $\dim \operatorname{im} \Omega = 10$. Then there do not exist $m_1, m_2, m_3 \in \mathbb{C}[x, y, z]_{\leq 1}$ such that $\Omega(m_1, m_2, m_3)$ contains terms of degree three. This implies that for all $i \in \{1, 2, 3\}$ holds that $\det M_0^{(i,3)} \in \mathbb{C}[x, y, z]_{\leq 1}$.

Let $K = \{Q \subseteq \mathbb{C}[x, y, z]_{\leq 1} : \dim Q = 3\}$. This is the set of three-dimensional subspaces of $\mathbb{C}[x, y, z]_{\leq 1}$. Let $U \in K$ be some three-dimensional subspace of $\mathbb{C}[x, y, z]_{\leq 1}$, then there are two cases: Either $1 \in U$, or $1 \notin U$.

If case $1 \in U$, then there exists $(u, g) \in \mathcal{AGL}_3(\mathbb{C})$ such that for all $P \in U$ holds that $(u, g)P \in \langle 1, x, y\rangle$. When the case $1 \notin U$ is considered, then there exists $(u, g) \in \mathcal{AGL}_3(\mathbb{C})$ such that for all $P \in U$ holds that $(u, g)P \in \langle x, y, z\rangle$. So there are two orbits of $\mathcal{AGL}_3(\mathbb{C})$ on K.

As $\det M_0^{(i,3)} \in \mathbb{C}[x, y, z]_{\leq 1}$ for $i \in \{1, 2, 3\}$, there is a three-dimensional subspace of $\mathbb{C}[x, y, z]_{\leq 1}$ containing each of these minors. Call this subspace $U \in K$, so that $\det M_0^{(i,3)} \in U$ for $i \in \{1, 2, 3\}$. Either U is in the orbit of $\langle x, y, z\rangle$, or in the orbit of $\langle 1, x, y\rangle$.

There exists $(u, g) \in \mathcal{AGL}_3(\mathbb{C})$ such that for all $\det M_0^{(i,3)} \in \mathbb{C}[x, y, z]_{\leq 1} \in U$ either $(u, g) \det M_0^{(i,3)} \in \langle 1, x, y\rangle$ or $(u, g) \det M_0^{(i,3)} \in \langle x, y, z\rangle$.

After applying the (u, g)-transformation, this means that in the first case every minor does not contain a term with monomial z. Therefore the term z^2 cannot be created. In the second case, every minor does not contain a term with monomial 1, so the constant term cannot be created. So in both cases it is not possible to map Ω to $\mathbb{C}[x, y, z]_{\leq 2}$.

For this reason, it is not possible to find a M_0 of type (1) such that $\operatorname{im} \Omega \subseteq \mathbb{C}[x, y, z]_{\leq 2}$.

Analysis where M_0 is of the type (2)

Suppose the matrix M_0 is of type (2) as discussed above.

Let $M_0 = \begin{pmatrix} n_{1,1} & n_{1,2} & n_{1,3} \\ 0 & 0 & n_{2,3} \\ 0 & 0 & n_{3,3} \end{pmatrix}$ and $M_1 = \begin{pmatrix} m_{1,1} & m_{1,2} & m_{1,3} \\ m_{2,1} & m_{2,2} & m_{2,3} \\ m_{3,1} & m_{3,2} & m_{3,3} \end{pmatrix}$. Let $M = M_0 + M_1$, so:

$$M = \begin{pmatrix} m_{1,1} + n_{1,1} & m_{1,2} + n_{1,2} & m_{1,3} + n_{1,3} \\ m_{2,1} & m_{2,2} & m_{2,3} + n_{2,3} \\ m_{3,1} & m_{3,2} & m_{3,3} + n_{3,3} \end{pmatrix}$$

In this notation, $n_{i,j} \in \mathbb{C}[\mathcal{V}]_{\leq 1}$ and $m_{i,j} \in \mathbb{C}[\mathcal{D}]_{=1}[\mathcal{V}]_{\leq 1}$.

Use that $\det M = p$ and that p is a polynomial with coefficients in $\mathcal{D}$. Therefore p does not contain terms from $\operatorname{span}(\mathcal{D} \cdot \mathcal{D})$.

Compute $\det M$ by applying minor expansion to the first row. Denote with $M^{(i,j)}$ the (i, j)-minor matrix of M. Use that $\det M^{(1,3)} = m_{2,1}m_{3,2} - m_{2,2}m_{3,1} \in \mathbb{C}[\mathcal{D}]_{=2}[\mathcal{V}]_{\leq 1}$, that

$\det M^{(1,1)}, \det M^{(1,2)} \in \mathbb{C}\left[\mathcal{D}\right]_{=1}\left[\mathcal{V}\right]_{\leq 1}$, and that for all i, j: $m_{i,j} \in \mathbb{C}\left[\mathcal{D}\right]_{=1}\left[\mathcal{V}\right]_{\leq 1}$.

$$\begin{aligned}
\det M &= (\underbrace{m_{1,1}}_{\in\mathbb{C}[\mathcal{D}]_{=1}[\mathcal{V}]_{\leq 1}} + n_{1,1})\underbrace{\det M^{(1,1)}}_{\mathbb{C}[\mathcal{D}]_{=1}[\mathcal{V}]_{\leq 1}} - (\underbrace{m_{1,2}}_{\in\mathbb{C}[\mathcal{D}]_{=1}[\mathcal{V}]_{\leq 1}} + n_{1,2})\underbrace{\det M^{(1,2)}}_{\in\mathbb{C}[\mathcal{D}]_{=1}[\mathcal{V}]_{\leq 1}} \\
&\quad + (m_{1,3} + n_{1,3})\underbrace{\det M^{(1,3)}}_{\in\mathbb{C}[\mathcal{D}]_{=2}[\mathcal{V}]_{\leq 1}} \\
&= \underbrace{n_{1,1}\det M^{(1,1)} - n_{1,2}\det M^{(1,2)}}_{\in\mathbb{C}[\mathcal{D}]_{=1}[\mathcal{V}]_{\leq 1}} \\
&\quad + \underbrace{m_{1,1}\det M^{(1,1)} - m_{1,2}\det M^{(1,2)} + (n_{1,3} + m_{1,3})\det M^{(1,3)}}_{\in\mathbb{C}[\mathcal{D}]_{=2}[\mathcal{V}]_{\leq 1}} \\
&= p \in \mathbb{C}\left[\mathcal{D}\right]_{=1}\left[\mathcal{V}\right]_{\leq 2}
\end{aligned}$$

From this computation it follows that $m_{1,1}\det M^{(1,1)} - m_{1,2}\det M^{(1,2)} + (n_{1,3} + m_{1,3})\det M^{(1,3)} = 0$, because these terms are element of $\mathbb{C}\left[\mathcal{D}\right]_{=2}\left[\mathcal{V}\right]_{\leq 1}$, which do not occur in p. Because these terms cancel out to each other, and do not occur anywhere else in the equation, it is possible to assume without loss of generality that $m_{1,1} = m_{1,2} = m_{1,3} = n_{1,3} = 0$.

What follows is a reduced form of the matrix M:

$$M = \begin{pmatrix} n_{1,1} & n_{1,2} & 0 \\ m_{2,1} & m_{2,2} & m_{2,3} + n_{2,3} \\ m_{3,1} & m_{3,2} & m_{3,3} + n_{3,3} \end{pmatrix}$$

Now compute $\det M$ by expanding the third column into minors.

$$\begin{aligned}
\det M &= -(\underbrace{m_{2,3}}_{\in\mathbb{C}[\mathcal{D}]_{=1}[\mathcal{V}]_{\leq 1}} + n_{2,3})\underbrace{\det M^{(2,3)}}_{\in\mathbb{C}[\mathcal{D}]_{=1}[\mathcal{V}]_{\leq 1}} + (\underbrace{m_{3,3}}_{\in\mathbb{C}[\mathcal{D}]_{=1}[\mathcal{V}]_{\leq 1}} + n_{3,3})\underbrace{\det M^{(3,3)}}_{\in\mathbb{C}[\mathcal{D}]_{=1}[\mathcal{V}]_{\leq 1}} \\
&= \underbrace{-n_{2,3}\det M^{(2,3)} + n_{3,3}\det M^{(3,3)}}_{\in\mathbb{C}[\mathcal{D}]_{=1}[\mathcal{V}]_{\leq 1}} + \underbrace{-m_{2,3}\det M^{(2,3)} + m_{3,3}\det M^{(3,3)}}_{\in\mathbb{C}[\mathcal{D}]_{=2}[\mathcal{V}]_{\leq 1}} \\
&= p \in \mathbb{C}\left[\mathcal{D}\right]_{=1}\left[\mathcal{V}\right]_{\leq 2}
\end{aligned}$$

By the same reasoning above, without loss of generality one can state that $m_{2,3} = m_{3,3} = 0$. The newly found reduced matrix M follows:

$$M = \begin{pmatrix} n_{1,1} & n_{1,2} & 0 \\ m_{2,1} & m_{2,2} & n_{2,3} \\ m_{3,1} & m_{3,2} & n_{3,3} \end{pmatrix}$$

In this matrix, for all i, j it holds that $n_{i,j}$ is a degree-one polynomial independent of $\mathcal{D}$ in the variables x, y and z, and that $m_{i,j}$ is a degree-one polynomial with coefficients depending on $\mathcal{D}$ in the variables x, y and z.

To prevent having too many indices, write M differently as follows:

$$M = \begin{pmatrix} a & b & 0 \\ m_1 & m_2 & c \\ m_3 & m_4 & d \end{pmatrix}$$

In this notation, $a, b, c, d \in \mathbb{C}[\mathcal{V}]_{\leq 1}$, and $m_1, m_2, m_3, m_4 \in \mathbb{C}[\mathcal{D}]_{=1}[\mathcal{V}]_{\leq 1}$.

Define $\Omega : \mathbb{C}[x,y,z]^4_{\leq 1} \to \mathbb{C}[x,y,z]_{\leq 3}$ as follows:

$$\Omega(m_1, m_2, m_3, m_4) = \det M = -m_1 bd + m_2 ad + m_3 bc - m_4 ac$$

Showing that $\mathbb{C}[x,y,z]_{\leq 2} \subseteq \operatorname{im}\Omega$ implies that the image of Ω contains every polynomial $p \in \mathbb{C}[x,y,z]_{\leq 2}$. Ω is represented by a determinant of a 3×3-matrix $M = M_0 + M_1$ with M_0 of type (2). This means that showing $\mathbb{C}[x,y,z]_{\leq 1} \subseteq \operatorname{im}\Omega$ corresponds to showing existence of a 3×3-matrix $M = M_0 + M_1$ with M_0 of type (2), of which the determinant can be any polynomial from $\mathbb{C}[x,y,z]_{\leq 2}$.

It holds that $\dim \mathbb{C}[x,y,z]_{\leq 1} = 4$. Therefore $\dim \mathbb{C}[x,y,z]^4_{\leq 1} = 16$. Therefore the domain of Ω has dimension 16.

By the assumption that $\operatorname{im}\Omega$ contains $\mathbb{C}[x,y,z]_{\leq 2}$, for the dimension of the image of Ω the following equality must hold: $\dim \operatorname{im}\Omega \geq \dim \mathbb{C}[x,y,z]_{\leq 2} = 10$.

Assume that $\dim\langle a, b\rangle_{\mathbb{C}} = \dim\langle c, d\rangle_{\mathbb{C}} = 2$. This is allowed, because assuming $\dim\langle a, b\rangle_{\mathbb{C}} \leq 1$ will cause a contradiction. This claim will be shown next.

Suppose that $\dim\langle a, b\rangle_{\mathbb{C}} \leq 1$ then there exists a $\lambda \in \mathbb{C}$ such that $b = \lambda a$. If this holds, then the following is found:

$$\Omega(m_1, m_2, m_3, m_4) = (m_2 - \lambda m_1)ad + (\lambda m_3 - m_4)ac = \Omega(0, m_2 - \lambda m_1, 0, \lambda m_3 - m_4)$$

Since $(m_2 - \lambda m_1), (\lambda m_3 - m_4) \in \mathbb{C}[x,y,z]_{\leq 1}$, it follows that $\dim \operatorname{im}\Omega \leq 2 \dim \mathbb{C}[x,y,z]_{\leq 1} = 2 \cdot 4 = 8$. This is in contradiction with $\operatorname{im}\Omega \supseteq \mathbb{C}[x,y,z]_{\leq 2}$, because then $\dim \operatorname{im}\Omega \geq \dim \mathbb{C}[x,y,z]_{\leq 2} = 10$.

So assuming a map Ω exists such that $\operatorname{im}\Omega \supseteq \mathbb{C}[x,y,z]_{\leq 2}$, it must hold that $\dim\langle a, b\rangle_{\mathbb{C}} = 2$. By symmetry it must also hold that $\dim\langle c, d\rangle_{\mathbb{C}} = 2$. So $\dim\langle a, b\rangle_{\mathbb{C}} = \dim\langle c, d\rangle_{\mathbb{C}} = 2$ must hold.

Because $\Omega(m_1, m_2, m_3, m_4) = -m_1 bd + m_2 ad + m_3 bc - m_4 ac$, the image of Ω can be defined as follows:

$$\operatorname{im}\Omega = \langle a, b\rangle_{\mathbb{C}} \cdot \langle c, d\rangle_{\mathbb{C}} \cdot \mathbb{C}[x,y,z]_{\leq 1}$$

Because $\dim\langle a, b\rangle_{\mathbb{C}} = \dim\langle c, d\rangle_{\mathbb{C}} = 2$, it holds that $\dim(\langle a, b\rangle_{\mathbb{C}} \cdot \langle c, d\rangle_{\mathbb{C}})$ is either 2, 3 or 4. These three cases will be considered next.

Case 1 When $\dim(\langle a, b\rangle_{\mathbb{C}} \cdot \langle c, d\rangle_{\mathbb{C}}) = 2$, then $\dim \operatorname{im}\Omega = \dim(\langle a, b\rangle_{\mathbb{C}} \cdot \langle c, d\rangle_{\mathbb{C}} \cdot \mathbb{C}[x,y,z]_{\leq 1}) \leq 2 \cdot 4 = 8$. If $\mathbb{C}[x,y,z]_{\leq 2} \subseteq \operatorname{im}\Omega$, it must hold that $\operatorname{im}\Omega$ is at least 10, so this is not possible $\lightning$.

Case 2 When $\dim(\langle a, b\rangle_{\mathbb{C}} \cdot \langle c, d\rangle_{\mathbb{C}}) = 3$, then $\dim \operatorname{im}\Omega = \dim(\langle a, b\rangle_{\mathbb{C}} \cdot \langle c, d\rangle_{\mathbb{C}} \cdot \mathbb{C}[x,y,z]_{\leq 1}) \leq 3 \cdot 4 = 12$. This case will be considered next.

Because $\dim\langle a, b\rangle_{\mathbb{C}} = \dim\langle c, d\rangle_{\mathbb{C}} = 2$, at least one of a and b is a nonconstant polynomial, and at least one of c and d is a nonconstant polynomial. Without loss of generality, suppose that b and d are nonconstant polynomials: $b, d \in \mathbb{C}[x,y,z]_{\leq 1}$. Define $B = \operatorname{im}\Omega \cap \mathbb{C}[x,y,z]_{\leq 2}$ and $A = \operatorname{im}\Omega / B$, such that $\dim \operatorname{im}\Omega = \dim A + \dim B$. The term bd is a term of degree two, so the term m_2bd is a term of degree three. Looking at the degree-three parts of the term m_2bd, it follows that $\mathrm{im}\{m_2bd : \deg m_2bd = 3\}$ has dimension 3 in A. What follows is a lower bound for the dimension of A: $\dim A \geq 3$.

Using this result one can estimate the dimension of B by using $\dim \mathrm{im}\,\Omega = \dim A + \dim B$:

$$\dim B = \dim \mathrm{im}\,\Omega - \dim A \leq 12 - 3 = 9$$

So $\dim B = \dim(\mathrm{im}\,\Omega \cap \mathbb{C}[x, y, z]_{\leq 2}) \leq 9$. Because $\dim \mathbb{C}[x, y, z]_{\leq 2} = 10$ and $\mathbb{C}[x, y, z]_{\leq 2} \subseteq \mathrm{im}\,\Omega$ should hold, this yields a contradiction $\lightning$.

Case 3 Lastly consider the case where $\dim(\langle a, b\rangle_{\mathbb{C}} \cdot \langle c, d\rangle_{\mathbb{C}}) = 4$.

Note that $\langle (a, b, 0, 0), (c, 0, d, 0), (0, c, 0, d), (0, 0, a, b)\rangle_{\mathbb{C}} \subseteq \ker \Omega$. This is easily verified by filling these entries in in the definition for the map Ω. Note that this does not necessarily imply that the kernel is at least four-dimensional, because one of the elements from the span can be a linear combination of the other elements. The claim is that this is not the case, and that therefore $\dim \ker \Omega \geq 4$. This claim will be shown next.

Because $1 \in \mathbb{C}[x, y, z]_{\leq 2}$, and because it is needed that $\mathbb{C}[x, y, z]_{\leq 2} \subseteq \mathrm{im}\,\Omega$, it is needed that $1 \in \mathrm{im}\,\Omega$ for the existence of the matrix. It holds that $\langle a, b\rangle_{\mathbb{C}}$, $\langle c, d\rangle_{\mathbb{C}}$ and $\mathbb{C}[x, y, z]_{\leq 1}$ are polynomial of maximal degree one, and that no negative degrees occur. For this reason, it must hold that $1 \in \langle a, b\rangle_{\mathbb{C}}$ and $1 \in \langle c, d\rangle_{\mathbb{C}}$. This is needed to satisfy the requirement $1 \in \mathrm{im}\,\Omega$. So without loss of generality, assume $a = c = 1$. Because $\dim\langle a, b\rangle_{\mathbb{C}} = \dim\langle a, b\rangle_{\mathbb{C}} = 2$ it can also be assumed that $b, d \in \mathbb{C}[x, y, z]_{=1}$.

Using $\dim(\langle a, b\rangle_{\mathbb{C}} \cdot \langle c, d\rangle_{\mathbb{C}}) = 4$ and $\langle a, b\rangle_{\mathbb{C}} \cdot \langle c, d\rangle_{\mathbb{C}} = \langle 1, b\rangle_{\mathbb{C}} \cdot \langle 1, d\rangle_{\mathbb{C}} = \langle 1, b, d, bd\rangle_{\mathbb{C}}$, it follows that b and d must be linearly independent, so $\dim\langle b, d\rangle_{\mathbb{C}} = 2$.

In the span $\langle (a, b, 0, 0), (c, 0, d, 0), (0, c, 0, d), (0, 0, a, b)\rangle_{\mathbb{C}} \subseteq \ker \Omega$, or in the simplified form $\langle (1, b, 0, 0), (1, 0, d, 0), (0, 1, 0, d), (0, 0, 1, b)\rangle_{\mathbb{C}}$, it is obvious that the first three entries of the span are independent, as they have zero-entries at different slots, and because $1 \in \mathbb{C}$ and $b, d \in \mathbb{C}[x, y, z]_{=1}$ are linearly independent. The fourth entry of the span could possibly be in the span of the others, so that $(0, 0, 1, b) \in \langle (1, b, 0, 0), (1, 0, d, 0), (0, 1, 0, d)\rangle_{\mathbb{C}}$. However, since b and d are linearly independent, and only the entries $(0, 1, 0, d)$ and $(0, 0, 1, b)$ have an entry in the fourth slot, this implies that it is not possible the write the last entry as linear combination of the first three entries.

Therefore, $\langle (1, b, 0, 0), (1, 0, d, 0), (0, 1, 0, d), (0, 0, 1, b)\rangle_{\mathbb{C}}$ is a four-dimensional linear span. Since this span is contained in $\ker \Omega$, it implies that $\dim \ker \Omega \geq 4$.

Use that $\Omega : \mathbb{C}[x, y, z]^4_{\leq 1} \to \mathbb{C}[x, y, z]$ and rank-nullity to find that $\dim \mathrm{im}\,\Omega + \dim \ker \Omega = \dim \mathbb{C}[x, y, z]^4_{\leq 1}$. Combining $\dim \ker \Omega \geq 4$ and $\dim \mathbb{C}[x, y, z]^4_{\leq 1} = 4 \dim \mathbb{C}[x, y, z]_{\leq 1} = 4 \cdot 4 = 16$, it follows that $\dim \mathrm{im}\,\Omega \leq 16 - 4 = 12$.

Now apply the same steps as done for $\dim(\langle a, b\rangle_{\mathbb{C}} \cdot \langle c, d\rangle_{\mathbb{C}}) = 3$. Let $B = \mathrm{im}\,\Omega \cap \mathbb{C}[x, y, z]_{\leq 2}$ and $A = \mathrm{im}\,\Omega/B$ such that $\dim A + \dim B = \dim \mathrm{im}\,\Omega$. Since $b, d \in \mathbb{C}[x, y, z]_{=1}$, it holds that $\deg bc = 2$. The term bd is a term of degree two, so the term m_2bd is a term of degree three. Looking at the degree-three parts of the term m_2bd, it follows that $\mathrm{im}\{m_2bd : \deg m_2bd = 3\}$ has dimension 3 in A. What follows is a lower bound for the dimension of A: $\dim A \geq 3$.

The same way: $\dim B = \dim \operatorname{im} \Omega - \dim A \leq 12 - 3 = 9$. Since $\dim B = \dim(\operatorname{im} \Omega \cap \mathbb{C}[x,y,z]_{\leq 2}) \leq 9$ and $\dim \mathbb{C}[x,y,z]_{\leq 2} = 10$, it follows that $\mathbb{C}[x,y,z]_{\leq 2} \not\subseteq \operatorname{im} \Omega$. This yields a contradiction $\lightning$.

Conclusion Since a contradiction follows for the cases where $\dim(\langle a,b\rangle_{\mathbb{C}} \cdot \langle c,d\rangle_{\mathbb{C}})$ is either 2, 3 or 4, and because these are all possible cases, it follows that it is not possible to find a 3×3-matrix $M = M_0 + M_1$ where M_0 is of type (2) such that the determinant of M is a general degree-two trivariate polynomial.

Analysis where M_0 is of the type (3)

Let $M_0 = \begin{pmatrix} 0 & n_{1,2} & n_{1,3} \\ -n_{1,2} & 0 & n_{2,3} \\ -n_{1,3} & -n_{2,3} & 0 \end{pmatrix}$ and $M_1 = \begin{pmatrix} m_{1,1} & m_{1,2} & m_{1,3} \\ m_{2,1} & m_{2,2} & m_{2,3} \\ m_{3,1} & m_{3,2} & m_{3,3} \end{pmatrix}$. Define $M = M_0 + M_1$ such that $\det M = p \in \mathbb{C}[x,y,z]_{\leq 2}$:

$$M = \begin{pmatrix} m_{1,1} & m_{1,2} + n_{1,2} & m_{1,3} + n_{1,3} \\ m_{2,1} - n_{1,2} & m_{2,2} & m_{2,3} + n_{2,3} \\ m_{3,1} - n_{1,3} & m_{3,2} - n_{2,3} & m_{3,3} \end{pmatrix}$$

Here, $n_{i,j} \in \mathbb{C}[\mathcal{V}]_{\leq 1}$ and $m_{i,j} \in \mathbb{C}[\mathcal{D}]_{=1}[\mathcal{V}]_{\leq 1}$.

Use that $\det M = p$ and p is a polynomial with coefficients in $\mathcal{D}$. Therefore p does not contain terms with coefficients from $\operatorname{span}(\mathcal{D} \cdot \mathcal{D})$. Therefore, the terms in $\det M$ that contains products of two or more entries from M_1 will cancel in $\det M$.

In other words, in the simplified expanded expression $\det M$ one finds that every term comes from exactly one entry from M_1, and exactly two entries from M_0. The entries from M_0 and M_1 are degree-one polynomials: $n_{i,j}, m_{i,j} \in \mathbb{C}[x,y,z]_{\leq 1}$. As there are three variables in p, this means that $n_{i,j} \in \langle 1, x, y, z\rangle$.

The claim that $n_{i,j} \in \langle 1, x, y, z\rangle$ for all i and j will be clarified next. Suppose that $n_{i,j} \in \langle x, y, z\rangle$ for all i and j. Each term of p is a term from the product of two entries from M_0 and one from M_1. If $n_{i,j} \in \langle x, y, z\rangle$, then it holds that all terms from p are of degree-two, because each term of $n_{i,j}$ is of degree one. But p could also contain terms of degree zero and one, so this is not possible.

Suppose that $n_{i,j} \in \langle 1, x, y\rangle$. Each term of p is a term from the product of two entries from M_0 and one entry from M_1. Since M_0 does not contain the variable z, there cannot be a term in p containing the monomial z^2, so this is not possible.

For this reason, the entries of M_0 must be degree-one polynomials in x, y and z. In other words: $n_{i,j} \in \langle 1, x, y, z\rangle$ for all i and j.

This is a necessary condition for the existence of a matrix M with skew-symmetric part independent of $\mathcal{D}$, such that that $\det M$ could represent every trivariate degree-two polynomial p. Next is to be shown that this condition cannot be satisfied for a skew-symmetric M_0.

For all $S, R \in \mathcal{SL}_3(\mathbb{C})$ it holds that $\det S = \det R = 1$. For this reason, $p = \det M = \det(SMR)$ for all $S, R \in \mathcal{SL}_3(\mathbb{C})$. Let $R = S^T$ and $S := S_{i,j} = I + \lambda \underline{e_i} \underline{e_j}^T$, so S is a matrix with ones on the diagonal and with an entry λ added at position (i,j). For example, $S_{1,2} = \begin{pmatrix} 1 & \lambda & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Clearly $S, S^T \in \mathcal{SL}_3(\mathbb{C})$ if $i \neq j$.

Computing $S^T M_0 S$ with $S = S_{i,j}$, then column i gets added λ times to column j ($M_0 S$), and row i is added λ times to column j. It is found that the matrix after this transformation is still skew-symmetric: $M_0^T = -M_0$, also $(S^T M_0 S)^T = S^T M_0^T S = -S^T M_0 S$.

Using the matrices $S_{1,2}$, $S_{1,3}$, $S_{2,1}$, $S_{2,3}$, $S_{3,1}$, $S_{3,2}$ on M_0, one can respectively add multiples of $n_{1,3}$ to $n_{2,3}$, $n_{1,2}$ to $n_{2,3}$, $n_{2,3}$ to $n_{1,3}$, $n_{1,2}$ to $n_{1,3}$, $n_{1,2}$ to $n_{2,3}$ and $n_{1,3}$ to $n_{1,2}$.

Let $K = \{Q \subseteq \mathbb{C}[x,y,z]_{\leq 1} : \dim Q = 3\}$, the set of three-dimensional subspaces of $\mathbb{C}[x,y,z]_{\leq 1}$. Define $U = \langle n_{1,2}, n_{1,3}, n_{2,3} \rangle$, then there is a $Q \in K$ such that $U \subseteq Q$ (and equality if $\dim U = 3$).

Because $\dim U \leq 3$ there exist $(u,g) \in \mathcal{AGL}_3(\mathbb{C})$ such that for all $P \in U$ either $(u,g)P$ contains a constant term (1), or $(u,g)P$ does not contain a constant term (2). In case (1), without loss of generality $(u,g)n_{i,j} \in \langle 1, x, y \rangle$. In case (2), without loss of generality $(u,g)n_{i,j} \in \langle x, y, z \rangle$.

By the reasoning above, both cases lead to a contradiction. For this reason, it is not possible to find a 3×3-matrix M of type (3) that has a general degree-two trivariate polynomial as determinant.

Conclusion for degree-two trivariate polynomials represented by 3×3-matrices

In the initial section about degree-two trivariate polynomials, it is stated that M_0 must be of the type (1), (2) or (3) if a 3×3-matrix representation exists for these polynomials. In the sections following to that, it has been shown that for all of these cases it is proven that there is no representation possible.

For this reason, it is impossible to find a 3×3 matrix representation of which the matrix entries are affine linear in the variables and affine linear in the coefficients of general degree-two trivariate polynomials, such that the determinant is the polynomial.

The construction of Plestenjak and Hochstenbach provided a matrix representation for degree-two trivariate polynomials of size 4×4. As 3×3 is impossible, this shows that their construction is minimal for degree-two trivariate polynomials.

Chapter 5

Conclusion and Outlook

In Chapter 3 a number of constructions are given. First the Frobenius Companion Matrix is discussed, which provides a matrix representation for monic univariate polynomials p of degree d in a $d \times d$-matrix that can be easily constructed. It is also shown that this is a minimal representation of which the entries of the matrix are affine linear in the coefficients and in the variable of p.

Next the construction proposed by Plestenjak and Hochstenbach [3] for bivariate polynomials of any degree is discussed, and this construction is generalized for more variables. Using this generalized construction, for every possible polynomial with complex-valued coefficients a matrix can be found of which the determinant is that polynomial. Considering a general polynomial with n variables and of degree d, it is found that the size of this matrix provided by this construction has a number of rows and columns that is of order $\frac{d^n}{2n!}$.

From this construction another is derived in which the exponent is considerably lower. This improved construction yields matrices of which the number of rows and columns is of order $\mathcal{O}(d^{\lceil \frac{n}{2} \rceil})$. The reduction of the exponent of the leading term by roughly a half is certainly a considerable improvement compared to generalized method. For this reason, if the degree is sufficiently high, the improved method yields matrices of smaller size than the generalized method.

Chapter 4 consists of two parts. In the first part a lower bound for the number of rows and columns of the matrix is derived for a polynomial with any number of variables and any degree. In the second part tight lower bounds are derived for matrices of which the determinant is a general polynomial of a certain degree and a certain number of variables.

Considering a general polynomial with n variables and of degree d, the best general lower bound that is found for the size of $N \times N$-matrices is $N \geq \sqrt{\frac{1}{n}\left(\binom{n+d}{d} - 1\right)}$. As the leading term of $\binom{n+d}{d}$ is $\frac{d^n}{n!}$, it follows that the order of this bound is $\frac{d^{n/2}}{\sqrt{n \cdot n!}}$. Note that the exponent $\frac{n}{2}$ of d does also occur in the improved method, so considering the order-term $d^{\frac{n}{2}}$, the improved method is close to this lower bound. However, when it comes to the factor $d^{\frac{n}{2}}$ is multiplied with, one can see that the size of the improved method is far from the theoretical lower bound. The improved method will have a factor that is at least $\frac{1}{\lceil \frac{n}{2} \rceil !}$ that is multiplied by $d^{\lceil \frac{n}{2} \rceil}$, whereas the lower bound provides a factor $\frac{1}{\sqrt{nn!}}$. Already for small values of n, $\sqrt{nn!}$ is significantly larger than $\lceil \frac{n}{2} \rceil!$. This means that the factor of the general lower bound is significantly smaller than the factor of the constructed upper bound.

The best construction introduced in this thesis is asymptotically to d of the same order as the theoretical general lower bound that is found. This means that asymptotically the theoretical lower bound is matched. However, there is a significant difference between the factor that the leading terms of the lower bound and construction are multiplied with. This means that there is still a lot of room for improvement. First of all, it is neither proven nor claimed that the derived general lower bound is tight, which means there may be room for improvement for the lower bound. Another point of improvement could be the construction of the improved method. In Section 3.5.2 a few proposals are made about how the improved method can yield matrices of smaller size, but these have not been worked out. This means that there may also be room for improvement when it comes to the construction of the matrices from the improved method.

In the second part of Chapter 4, tight lower bounds for the multivariate one-dimensional, the bivariate two-dimensional and the trivariate two-dimensional general polynomial are found.

According to the general lower bound $N \geq \sqrt{\frac{1}{n}\left(\binom{n+d}{d} - 1\right)}$, the matrices whose determinant is a multivariate degree-one polynomial should be of size 1×1. For bivariate degree-two polynomials this should be $N \geq \sqrt{5/2}$, so the matrix size should be at least 2×2. For the bivariate degree-three this should be $N \geq \frac{3}{\sqrt{2}}$, so the size should be at least 3×3.

The multivariate one-dimensional case is a silly case, because this polynomial can be taken as entry in an 1×1 matrix, which will have the entry as determinant. So what follows is that the general lower bound yields a tight bound for this specific case.

The other two cases are less trivial. In both the degree-two bivariate case as well as the degree-two trivariate case, it has been shown that the construction of Plestenjak and Hochstenbach provides matrices of minimal size, because contradictions follow if one searches for matrices of smaller size. This means that in the degree-two bivariate case, a matrix representation is at least of size 3×3, and in the case of degree-two trivariate polynomials a matrix representation is at least of size 4×4. Note that the general lower bound came up with respectively 2×2 and 3×3. This shows that the derived general lower bound is not a tight lower bound.

As stated, there is still a lot of room for improvement of the research done in this thesis. As the general lower bound for matrix sizes $N \geq \sqrt{\frac{1}{n}\left(\binom{n+d}{d} - 1\right)}$ is not tight, this lower bound can be improved. Also there may be determinantal representations that provides matrices of smaller size.

Bibliography

[1] S. Neumark, *Solution of Cubic and Quartic Equations.* Elsevier Science, 2014.

[2] J. Brown, "Abel and the insolvability of the quintic," 2005. [Online]. Available: http://www.math.caltech.edu/~jimlb/abel.pdf

[3] B. Plestenjak and M. E. Hochstenbach, "Roots of bivariate polynomial systems via determinantal representations." *SIAM J. Sci. Comput.*, vol. 38, no. 2, pp. A765–A788, 2016.

[4] X.-M. Niu and T. Sakurai, "A method for finding the zeros of polynomials using a companion matrix," *Japan journal of industrial and applied mathematics*, vol. 20, no. 2, pp. 239–256, 2003. [Online]. Available: http://link.springer.com/article/10.1007%2FBF03170428

[5] J. van Lint and J. van Nienhuys, *Discrete Wiskunde.* Academic Service, Schoonhoven, 1991, no. ISBN 90 6233 368 0, pp. 3–8.

[6] P. Fillmore, C. Laurie, and H. Radjavi, "On matrix spaces with zero determinant," *Linear and Multilinear Algebra*, vol. 18, no. 3, pp. 255–266, 1985.