Computing Valuation Popov Forms

Mark Giesbrecht[1], George Labahn[1], and Yang Zhang[2]

[1] School of Computer Science, University of Waterloo,
Waterloo, ON, N2L 3G1, Canada
{`mwg, glabahn`}`@uwaterloo.ca`
[2] Dept. of Mathematics and Computer Science, Brandon University,
Brandon, MB, R7A 6A9, Canada
`zhangy@brandonu.ca`

Abstract. Popov forms and weak Popov forms of matrices over non-commutative valuation domains are defined and discussed. Two new algorithms to construct these Popov forms are given, along with a description of some of their applications.

1 Introduction

Over the last decade, there have been significant advances in computing with symbolic systems of linear recurrence and differential equations. At the core of this new technology are computations with Ore polynomials, which provide a unified algebraic structure for these different domains. The Ore polynomials form a noncommutative polynomial ring which acts on a space of functions as a differential or difference (recurrence) operator. Matrices of such polynomials describe linear differential and differential systems of equations, which are clearly an important topic in computer algebra systems. Casting differential and recurrence operators as polynomial-like objects allows us to potentially employ the now very well developed and very efficient algorithms for matrices of polynomials in this new setting. This paper is a step towards efficient algorithms for such systems.

It is well known that the Hermite normal form of a polynomial matrix has degree bounds that can result in large row or column degree values. While this is not a disadvantage for some applications (such as linear solving) it can be a significant problem in such computations as matrix-polynomial GCDs. A second normal form, now known as the Popov normal form, was introduced by Popov [11] and has the property that row degrees are controlled. These normal forms have been successfully applied in such applications as control theory and linear systems (see for example, Kailath[8]).

Recently the Popov normal form and some variations (the shifted Popov form and the weak Popov form) have attracted considerable interest in the computer algebra community, for example, in Villard [13], Beckermann, Labahn and Villard [4], and Mulders and Storjohann [9]. While the normal forms previously mentioned all deal with matrices of commutative polynomials, Popov

normal forms and their variations also exist for some matrices of noncommutative polynomials, particularly Ore polynomials. Ore domains were introduced in the 1930's as a unified way of capturing differential and recurrence operators (c.f., Ore [10], Jacobson [7]). In 1943 Jacobson [7] considered the Hermite and Smith forms of matrices over skew polynomial rings. In the computer algebra area, Abramov and Bronstein [1] gave a method to calculate the ranks of matrices over skew polynomial rings using a type of normal form reduction of these matrices. Recently Beckermann, Cheng and Labahn [3] used fraction-free methods to compute weak Popov form of matrices over skew polynomial rings.

In this paper, we consider the more general problem of defining and computing Popov and weak Popov forms for matrices over *valuation rings*. The motivation for this question comes from several directions: Mulders and Storjohann [9] extended the notion of weak Popov forms to the discrete valuation rings, but failed to define Popov forms. Meanwhile, as they point out, their methods have the problem that they do not necessarily terminate. We give a definition of Popov forms of matrices over valuation Ore domains, and provide an algorithm for such forms which is guaranteed to terminate. Our methods lead to a unified treatment of many polynomial rings currently studied in Computer Algebra, for example, commutative polynomial rings ([13], [4] and [9]), skew polynomial rings ([3] and [1]) and multivariate polynomials ([6] and example 4). Finally, we remark that our method can be used to extend the work on reduced bases for matrices over commutative valuations rings by von zur Gathen [14] to some noncommutative cases ([6]).

It is worth noting that in many cases valuation methods are quite different from term ordering methods, which are the usual methods to deal with polynomials. When using term ordering methods, one only considers the leading terms at each reduction step. Valuation methods usually consider the whole polynomial. In fact, any term ordering induces a valuation but the converse is not true.

2   Valuations on Ore Domains

Noncommutative valuation theory dates back to the 1930's with the work of Artin, Noether and others. Presentations can be found in Schilling [12] and later in Artin [2], with a more recent treatment given in Cohn [5]. In this paper, we restrict such noncommutative valuations to the more practically motivated Ore domain case.

Let us first recall some definitions and discuss some properties of valuations on skew fields. Some results have already been discussed in different settings for commutative cases.

Definition 1. *An ordered group Γ is a (not necessarily commutative) group with a total ordering $\alpha \geq \beta$, which is preserved by the group operation:*

$$\alpha \geq \beta, \ \alpha' \geq \beta' \Rightarrow \alpha + \alpha' \geq \beta + \beta' \quad \text{for all } \alpha, \alpha', \beta, \beta' \in \Gamma.$$

Usually, Γ will be augmented by a symbol ∞ to form a monoid with the operation

$$\alpha + \infty = \infty + \alpha = \infty + \infty = \infty \quad \text{for all } \alpha \in \Gamma,$$

and the ordering $\infty > \alpha$ for all $\alpha \in \Gamma$.

Definition 2. *Let R be a ring. Then a valuation on R with values in an ordered group Γ (the value group) is a function ν on R with values in $\Gamma \cup \{\infty\}$ which satisfies:*

(V.1) $\nu(a) \in \Gamma \cup \{\infty\}$ and ν assumes at least two values,
(V.2) $\nu(ab) = \nu(a) + \nu(b)$, and
(V.3) $\nu(a + b) \geq min\{\nu(a), \nu(b)\}$, for every pair of elements $a, b \in R$.

Lemma 1. *([5]) A valuation ν on a ring R has the following properties:*

(a) $\nu(0) = \infty$ so that $\nu(0) > \gamma$ for all $\gamma \in \Gamma$ (by (V.2)).
(b) $ker\ \nu = \{a \in R \mid \nu(a) = \infty\}$ is a proper ideal of R (by (V.1)). Therefore, $ker\ \nu = 0$ if R is a skew field or a simple ring. Furthermore, $R/ker\ \nu$ is an integral domain.
(c) $\nu(1) = 0$ and $\nu(a) = \nu(-a)$.
(d) If $\nu(a) \neq \nu(b)$, then $\nu(a + b) = min\{\nu(a), \nu(b)\}$.   □

Remark 1. Note that $ker\ \nu$ is different from the set $\{a \in R \mid \nu(a) = 0\}$.

   In the case of a skew field K with valuation ν, there are two important properties of the valuation ring $V := \{a \in K \mid \nu(a) \geq 0\}$. First, the valuation ring is *total*: for every $k \in K$, either $k \in V$, or $k \neq 0$ and $k^{-1} \in V$. Second, the valuation ring is *invariant*: for every $a \in V$ and $k \in K \setminus \{0\}, k^{-1}ak \in V$. Usually a total invariant subring of a skew field is called a *valuation ring*. Conversely, given any skew field K and a valuation ring V in K, one can form a valuation which gives rise to V. Moreover we also have the following properties:

Lemma 2. *With our previous notation we have:*

(a) Given $a, b \in V$, then a is a left (and right) multiple of b if and only if $\nu(a) \geq \nu(b)$.
(b) The units set of V is $U = \{a \in V \mid \nu(a) = 0\}$. Furthermore, the non-units of V is $M = \{a \in V \mid \nu(a) > 0\}$, a maximal ideal of V.
   (i) $\nu(a) \geq \nu(b) \Leftrightarrow a = rb = br'$ for some $r, r' \in V$.
   (ii) $\nu(a) = \nu(b) \Leftrightarrow a = rb = br'$ for some $r, r' \in U$.
   (iii) $\nu(a) > \nu(b) \Leftrightarrow a = rb = br'$ for some $r, r' \in M$.

Proof.

(a) Assume that for some $r, s \in V$, we have $a = br = sb$. Then $\nu(a) = \nu(b) + \nu(r) = \nu(s) + \nu(b) \geq \nu(b)$ follows from the definition, $\nu(s) \geq 0$ and $\nu(r) \geq 0$. Conversely, if $b = 0$, then $\nu(b) = \infty$ implies $\nu(a) = \infty$. Therefore $a = 0$ so that a is both a left and right multiple of b. Now we can assume that $\nu(a) \geq \nu(b), b \neq 0$. This implies that $\nu(ab^{-1}) \geq 0, \nu(b^{-1}a) \geq 0$, that is, both $ab^{-1} \in V$ and $b^{-1}a \in V$. Therefore $a = (ab^{-1})b = b(b^{-1}a)$.

(b) Part (i) follows by (a). For (ii) it is easy to see that $\nu(r) = 0$ by the same proof as in (a). Therefore $r \in U$ and part (iii) follows from the definition of M. □

Example 1. (Cohn [5]) Let K be a skew field with a valuation ν and an automorphism σ of K such that $\nu(a^\sigma) = \nu(a)$ for all $a \in K$. Let γ be in the value group Γ (or in an ordered extension of Γ) and define a valuation on the skew polynomial $K[x; \sigma]$ by the rule

$$w(\sum a_i x^i) = min_i\{i\gamma + \nu(a_i)\}.$$

Then w is a valuation on $K[x; \sigma]$ and can be uniquely extended to a valuation of the function field $K(x; \sigma)$. If the residue-class skew field of K under ν is $\bar{K}$ and the automorphism induced on $\bar{K}$ by σ is $\bar{\sigma}$, then the residue-class field of $K(x; \sigma)$ is either $\bar{K}(x; \sigma^{-j})$ (when $j\gamma$ is the least multiple of γ which lies in Γ, that is, $j = 1$ if $\gamma \in \Gamma$ or there may exist a j such that $j\gamma \in \Gamma$ if $\gamma \notin \Gamma$) or $\bar{K}$ (when no multiple of γ lies in Γ).

The following lemma is well-known in the commutative case and gives the division rule in the sense of valuations. Now we extend it to the noncommutative case.

Lemma 3. *Let ν be a valuation on a skew field K with the valuation ring V, S be a subskew field of K that lies in V and $M = \{a \in V \mid \nu(a) > 0\}$. Assume that S maps isomorphically to the residue class skew field of ν, or equivalently, that $V = S \oplus M$ as abelian group. If $0 \neq a, b \in K$ such that $\nu(a) \geq \nu(b)$, then there exists a unique $s \in S$ such that either $a = sb$ or $\nu(a - sb) > \nu(b)$. Equivalently, there exists a unique $s \in S$ such that $a - sb = mb$ for some $m \in M$.*

Proof. From Lemma 2, $\nu(a) \geq \nu(b)$ implies that $a = rb$ for some $r \in V$. Since $V = S \oplus M$ there exists a unique $s \in S$ such that $r = s + m$ for some $m \in M$. Thus $a = rb = (s + m)b = sb + mb$, and so $a - sb = mb$. Clearly, $m = 0$ if and only if $a = sb$. If $m \neq 0$, then by Lemma 2 we have $\nu(mb) > \nu(b)$ and so $\nu(a - sb) > \nu(b)$. □

Definition 3. *Let ν be a valuation on a skew field K with the valuation ring V. For $a, b \in V$, we say that b right divides a with respect to ν, denoted $b \mid_\nu a$, if there exists $d \in V$ such that $\nu(a) = \nu(db)$. The element d is called a right valuation quotient of a by b with respect to ν if $a = db$, or if $a \neq db$ and $\nu(a - db) > \nu(a)$.*

It is easy to prove the following lemma.

Lemma 4. *$b \mid_\nu a$ if and only if there exists a right valuation quotient of a by b.*

In the remainder of this paper we will assume that ν is a valuation on a skew field K with the valuation ring V.

3   Valuation Popov Forms

In this section we define Popov and weak Popov forms for matrices over valuation domains. Unlike some papers discussing (weak) Popov forms, we consider the row vectors as a left module over a base ring and define the Popov form by two kinds of reductions. In this way the termination properties of reductions can be easily obtained, and one can use previously existing algorithms to compute these forms.

Definition 4. *Let $\boldsymbol{v} = (v_1, \cdots, v_m) \in V^m$. We define the pivot element $piv(\boldsymbol{v}) \in \{v_1, \cdots, v_m\}$ as the rightmost element with minimum valuation in $\boldsymbol{v}$.*

Next we define two kinds of reductions which are used for constructing the weak Popov form and the Popov form.

Definition 5. *Given $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \in V^m$, we say that*

(a) $\boldsymbol{a}$ weakly reduces to $\boldsymbol{c}$ modulo $\boldsymbol{b}$ in one step with respect to ν if and only if $piv(\boldsymbol{b})$ and $piv(\boldsymbol{a})$ have the same index, $piv(\boldsymbol{b}) \mid_\nu piv(\boldsymbol{a})$ and $\boldsymbol{c} = \boldsymbol{a} - q_1\boldsymbol{b}$, where $q_1 \in V$ is a valuation quotient of $piv(\boldsymbol{a})$ by $piv(\boldsymbol{b})$.

(b) $\boldsymbol{a}$ reduces to $\boldsymbol{c}$ modulo $\boldsymbol{b}$ in one step with respect to ν if and only if $piv(\boldsymbol{b}) \mid_\nu d$, where d is an entry that appears in $\boldsymbol{a}$ with the same index of $piv(\boldsymbol{b})$ and $\boldsymbol{c} = \boldsymbol{a} - q_2\boldsymbol{b}$, where $q_2 \in V$ is a valuation quotient of d by $piv(\boldsymbol{b})$.

Definition 6. *A nonzero vector $\boldsymbol{a}$ in V^m is (resp. weakly) reduced with respect to a set $S = \{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_l\}$ of nonzero vectors in V^m if $\boldsymbol{a}$ cannot be (resp. weakly) reduced modulo any one of $piv(\boldsymbol{s}_i)$ with respect to ν, $i = 1, \ldots, l$.*

Furthermore, a set $S = \{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_l\}$ of vectors is called a (resp. weakly) reduced set if any vector $\boldsymbol{s}_i$ is (weakly) reduced with respect to $S \setminus \{\boldsymbol{s}_i\}$.

Definition 7. *Let $\gamma = \{\boldsymbol{r}_1, \ldots, \boldsymbol{r}_n\}$ be the set of row vectors of matrix $V^{n \times m}$.*

(a) $V^{n \times m}$ is called in weak Popov form if γ is a weakly reduced set.

(b) $V^{n \times m}$ is called in Popov form if

    (i) γ is a reduced set.

    (ii) Rows are in an ascending chain with respect to ν, that is, $\nu(piv(\boldsymbol{r}_1)) \leq \nu(piv(\boldsymbol{r}_2)) \leq \cdots \leq \nu(piv(\boldsymbol{r}_n))$.

Note that our definition is the same as the usual one if we choose V as a commutative polynomial ring with degree valuation. A Popov form is a weak Popov form.

The following algorithm gives a method to construct a weak Popov form and also a Popov form.

Algorithm: `(resp. Weak) Popov form for` $V^{n \times m}$

Input: ▸ Row vectors $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_n$ of matrix $A \in V^{n \times m}$;
Output: ▸ Row vectors $\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n$ of a (resp. weak) Popov form of $A \in V^{n \times m}$;
       Initialization: $\boldsymbol{p}_1 := 0, \ldots, \boldsymbol{p}_n := 0$;

(1) Check if $\{\boldsymbol{r}_1, \ldots, \boldsymbol{r}_n\}$ is in a (resp. weak) Popov form;
(2)   If not
(3)   Swap rows of A to make them into a ascending chain w.r.t. ν;
(4)   For $i = 2$ to $i = n$ do
(5)   If $\boldsymbol{r}_i$ is (resp. weakly) reducible modulo $\boldsymbol{r}_1$,
(6) $\boldsymbol{r}_i := \boldsymbol{r}_i - q_i \boldsymbol{r}_1$, where q_i is the valuation quotient;
(8)   end if
(9)   goto (1);
(10)   else return: $\boldsymbol{p}_1 := \boldsymbol{r}_1, \ldots, \boldsymbol{p}_n := \boldsymbol{r}_n$;

Theorem 1. *If $\nu(V)$ is well-defined, then the preceding two algorithms are correct.*

Proof. We only need to prove the algorithms terminate. Note that every reduction increases the valuation, and forms an ascending chain. Therefore the algorithms will terminate after a finite number of steps since $\nu(V)$ is well-defined. □

The following gives a second method to construct (resp. weak) Popov forms:

Algorithm: `Constructing (resp. weak) Popov form for` $V^{n \times m}$

Input: ▸ row vectors $\gamma := \{\boldsymbol{r}_1, \ldots, \boldsymbol{r}_n\}$ of a matrix $A \in V^{n \times m}$;
Output: ▸ row vectors $\rho := \{\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n\}$ of (resp. weak) Popov form of $A \in V^{n \times m}$;

(1) Initialization: $\boldsymbol{p}_1 := 0, \ldots, \boldsymbol{p}_n := 0$;
(2) For $i = 1$ to $i = n$ do
(3) $\boldsymbol{p}_i := \boldsymbol{r}_i$ (resp. weakly) reduced modulo $\gamma := \{\gamma \setminus \{\boldsymbol{r}_i\}\} \cup \{\boldsymbol{p}_i\}$;
     end do;
(4) Swap rows such that they are in an ascending chain with respect to ν if Popov form required;

Theorem 2. *If $\nu(V)$ is well-defined, then the above algorithm is correct.*

Proof. The above algorithm terminates after a finite number of steps since $\nu(V)$ being well-defined implies that the steps in (3) terminate.
     We only need to check that ρ is a (weakly) reduced set with respect to ν. The first two steps produce $\rho = \{\boldsymbol{p}_1, \boldsymbol{r}_2, \cdots, \boldsymbol{r}_n\}$ and $\rho = \{\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{r}_3, \cdots, \boldsymbol{r}_n\}$. Therefore we have to prove $\boldsymbol{p}_1$ cannot be (weakly) reduced modulo $\boldsymbol{p}_2$. If this was the case, then $\boldsymbol{p}_1$ could be (weakly) reduced modulo $\boldsymbol{r}_2$, a contradiction. By induction, it is easy to prove that ρ is a (weakly) reduced set. □

4   Applications and Examples of Valuation Popov Forms

In the rest of this paper we mention some applications of valuation Popov forms. First, recall that the definition of the rank of a matrix over an Ore domain is different from the usual definition of rank (for example, see [5]). Given a matrix $A \in V^{n \times m}$, let M be the row space of A and M' be the reduced form with respect to (weak) reduction. From the previous algorithms, M and M' generate the same row space. Therefore we have the following.

Proposition 1. *The rank of M generated by the row space of $A \in V^{n \times m}$ is invariant under (weak) reduction with respect to any valuation.*

Proposition 2. *The rank of a (weak) Popov form equals its number of nonzero rows.*

Proof. This follows from the row set of (weak) Popov form being a (weakly) reduced set.   □

Example 2. The following example given in [9] is used to show that their procedures do not necessarily terminate. Using our algorithms one can easily obtain a Popov form as follows.

Let A be a matrix over a power series $F[[x]]$, where F is a field, that is,

$$A = \begin{bmatrix} x + x^2 + x^3 + \cdots \\ 1 \end{bmatrix} \in F[[x]]^{2 \times 1}.$$

Assume that ν is the usual valuation on the power series domain $F[[x]]$. Then $\nu(x + x^2 + x^3 + \cdots) = 1$ and $\nu(1) = 0$. Therefore the valuation quotient is $x + x^2 + x^3 + \cdots$. Note that $\nu(0) = \infty$. Our algorithm then computes the Popov form as $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

Example 3. Different valuations on polynomial rings may induce different Popov forms. For example, let R be a ring with a discrete valuation function $\nu(x)$. To extend the valuation $\nu(x)$ from R to the polynomial ring $R[t]$ we define, for an element $f[t] = a_0 + a_1 t + \cdots + a_n t^n$,

$$\nu(f[t]) = \min_i \{\nu(a_i) + i\} \tag{1}$$

It is well-known that ν defines a valuation function on $R[t]$ which extends $\nu(x)$. The algorithm for valuation Popov forms produces a Popov form which is different from the ones discussed in [4], [8] and [9].

Example 4. In some cases our algorithms can be used for multivariate domains. For example, let L be a finite dimensional Lie algebra over a commutative field K and let $U(L)$ be its universal envelope. It is well-known that $U(L)$ has a filtration

$$U^{-1} = 0; \ U^i = K + L + L^2 + \cdots + L^i$$

whose associated graded algebra is a polynomial algebra. This filtration then defines a valuation function $\nu(x)$ by the following rule:

$$\nu(0) = \infty, \quad \nu(x) = -i \text{ if } x \in U^i \setminus U^{i-1}. \tag{2}$$

Therefore, given any matrix over $U(L)$, we can use our algorithms to get Popov and weak Popov forms, and hence also easily obtain ranks.

5   Conclusion

In this paper we discuss some properties of valuations in noncommutative domains and define reductions for vectors in terms of valuations. For matrices over Ore domains we define and describe Popov and weak Popov forms in terms of valuations. Algorithms to construct these Popov forms are given and are shown to terminate. Further discussion and properties of these forms will appear in a forthcoming paper [6].

Acknowledgements

References

1. Abramov, S., Bronstein, M.: Linear Algebra for skew-polynomial matrices. INRIA (preprint)
2. Artin, E.: Geometric Algebra. Interscience Publishers, Inc. 1957
3. Beckermann, B., Cheng, H., Labahn, G.: Fraction-free row reduction of matrices of Ore polynomials. Proceedings of ISSAC'02, ACM Press, New York, 8-15
4. Beckermann, B., Labahn, G., Villard, G.: Shifted Normal Forms of Polynomial Matrices. Proceedings of ISSAC'99, ACM Press, ACM Press, New York, 189-196
5. Cohn, P. M.: Skew Fields. Cambridge University Press, 1995
6. Giesbrecht, M., Labahn, G., Zhang, Y.: Popov forms for multivariate polynomial matrices. (Preprint)
7. Jacobson, N.: The Theory of Rings. American Math. Soc., New York, 1943
8. Kailath, T.: Linear Systems. Prentice Hall, 1980
9. Mulders, T., Storjohann, A.: On lattice reduction for polynomial matrices. Journal of Symbolic Computation 35(2003), 377-401
10. Ore, O.: Theory of non-commutative polynomials. Annals of Mathematics 34(22): 480-508, 1933
11. Popov, V.: Invariant description of linear, time-invariant controllable systems. SIAM J. Control. Vol.10(2), 252-264, 1972
12. Schilling, O.: The Theory of Valuations. Math. Surveys No. IV, Amer. Math. Soc., New York 1950
13. Villard, G.: Computing Popov and Hermite forms of polynomial matrices. Proceeding of ISSAC'96, ACM Press, New York, 250-258
14. von zur Gathen, J.: Hensel and Newton methods in valuation rings. Mathematics of Computation vol.42(1984), 637-661