Efficient Computation of Nullspace Bases

May 20, 2011

1 Nullspace Basis Computation

1.1 High level outline

Let us first assume that the dimension of the $m \times n$ input matrix $\mathbf{F}$ satisfies $n = 3m$. Let $\vec{s} = [s_1, \ldots, s_n] \in \mathbb{Z}$ be a list of shift, such that each column degree of $\mathbf{F}$ is bounded by the corresponding entry in $\vec{s}$. For example, we can set each entry of $\vec{s}$ to be the corresponding column degree of $\mathbf{F}$, or we can simply set each entry of $\vec{s}$ to be the maximum column degree of $\mathbf{F}$. This condition on the shift is very useful, since it makes the column degrees of $\mathbf{FA}$ for any polynomial matrix $\mathbf{A}$ bounded component-wise by the $\vec{s}$-column degrees of $\mathbf{A}$. For simplicity, we also assume without loss of generality that the columns of $\mathbf{F}$ and the corresponding entries of $\vec{s}$ are arranged to make the entries of $\vec{s}$ in increasing orders. Let $\xi = \sum \vec{s}$. We want to compute a $\vec{s}$-minimal null space basis $\mathbf{N}$ with a cost of $h(n, \xi) = O^{\sim}(n^{\omega-1}\xi)$.

Let $s = \xi/n$ be the average of the entries of $\vec{s}$. First, we compute an $(\mathbf{F}, 4s, \vec{s})$-basis $\mathbf{P}$, which can be done efficiently. We claim that $\mathbf{P}$ contains at least $2n/5$ null space basis elements. (See Lemma 2 for a proof.)

Let $[\mathbf{P}_1, \mathbf{P}_2] = \mathbf{P}$ with $\mathbf{P}_1$ consists of the nullspace basis elements computed. Then the residual $\mathbf{FP} = [\mathbf{0}, \mathbf{FP}_2]$ can be used to compute the remaining nullspace basis elements. We claim that the matrix multiplication $\mathbf{FP}_2$ can be done efficiently. (See Corollary 1) If $\vec{b}$ is the $\vec{s}$-column degrees of $\mathbf{P}_2$ and $\mathbf{Q}$ is a $\vec{b}$-minimal nullspace basis of $\mathbf{FP}_2$, then $[\mathbf{P}_1, \mathbf{P}_2\mathbf{Q}]$ is a $\vec{s}$-minimal nullspace basis of $\mathbf{F}$.

Assume the columns of $\mathbf{P}$ are arranged in increasing order of their $\vec{s}$-column degrees. Then by Lemma 3 the $\vec{s}$-column degrees of $\mathbf{P}$ are bounded component-wise by $\vec{s} + 4s$. Let $\vec{t} = \vec{b} - 4s$, where $\vec{b}$ is the $\vec{s}$-column degrees of $\mathbf{P}_2$, then $\vec{t}$ is bounded component-wise by the largest entries of $\vec{s}$, with the sum of the entries bounded by ξ.

At this point, the problem is reduced to computing a $\vec{t}$-minimal nullspace basis of $\mathbf{G} = \mathbf{FP}_2/x^{4s}$, which still has row dimension m, but column dimension bounded by $3n/5$ and column degrees bounded by $\vec{t}$, with $\sum \vec{t}$ bounded by ξ.

Let

$$\begin{bmatrix} \mathbf{G}_1 \\ \mathbf{G}_2 \end{bmatrix} = \mathbf{G}$$

with $\mathbf{G}_1$having $n/5$ rows and $\mathbf{G}_2$ having $2n/15$ rows. If we compute a $\vec{t}$-minimal nullspace basis $\mathbf{N}_1$ of $\mathbf{G}_1$, where $\mathbf{N}_1$ has $\vec{t}$-degrees $\vec{u}$, then compute a $\vec{u}$-minimal nullspace basis $\mathbf{N}_2$ of $\mathbf{G}_2\mathbf{N}_1$, then $\mathbf{N}_1\mathbf{N}_2$ is a $\vec{t}$-minimal nullspace basis of $\mathbf{G}$. We claim that the sum of column degrees of $\mathbf{G}_2\mathbf{N}_1$ is bounded by ξ (Lemma 1), and the multiplication $\mathbf{G}_2\mathbf{N}_1$ can be computed efficiently (Corollary 2), and the multiplication $\mathbf{N}_1\mathbf{N}_2$ can also be done efficiently (Corollary 2).

The computation of $\mathbf{N}_1$ and $\mathbf{N}_2$ is similar to the original problem, only the dimension has decreased. For computing $\mathbf{N}_1$, the dimension of the input matrix $\mathbf{G}_1$ is bounded by $(n/5) \times (3n/5)$. For computing $\mathbf{N}_2$, and the dimension of input matrix $\mathbf{G}_2\mathbf{N}_1$ is bounded by$(2n/15) \times (2n/5)$. The sum of column degrees of $\mathbf{G}_1$ and that of $\mathbf{G}_2\mathbf{N}_1$ are both bounded by ξ still.

Let $h(n, \xi)$ be the computational cost of the original problem, then we have the recurrence relation

$$h(n, \xi) \in O^{\sim}(n^{\omega-1}\xi) + h(3n/5, \xi) + h(2n/5, \xi),$$

which gives $h(n, \xi) \in O^{\sim}(n^{\omega-1}\xi)$.

We now have an efficient way to compute a $\vec{s}$-minimal nullspace basis for input matrix with dimension $n = 3m$. For more general cases, if $n < 3m$ and $m \in \Theta(n)$, we can simply add zero columns to the input matrix to make $n = 3m$. If $n > 3m$, we can do a precomputation like in the first step to compute a partial nullspace basis that reduces the column dimension to $3m$. In this case, we can simply take ξ as the sum of the column degrees of the $3m$ columns with the largest column degrees.

In summary, this gives a way to compute a $\vec{s}$-minimal nullspace basis of $\mathbf{F}$ with a cost of $O^{\sim}(n^{\omega-1}\xi)$, where n is the column dimension of $\mathbf{F}$ and ξ is the sum of t he column degrees of the $3m$ columns with the largest column degrees. Of course, a small simplification for the computational complexity is that we can just take the sum λ of just m instead of $3m$ columns with the largest column degrees, since $O^{\sim}(n^{\omega-1}\xi) = O^{\sim}(n^{\omega-1}\lambda)$.

1.2 Some proofs

Now we give the proofs of some of the claims:

Lemma 1. *Given an $\mathbf{F} \in \mathbb{K}[x]^{m\times n}$, and a shift $\vec{s} \in \mathbb{Z}_{\geq 0}$ with entries bounding the corresponding column degrees of $\mathbf{F}$. The sum of the $\vec{s}$-column degrees of any $\vec{s}$-minimal nullspace basis of $\mathbf{F}$ is bounded by $\xi = \sum \vec{s}$.*

Proof. Let $\mathbf{P}$ be a $(\mathbf{F}, \vec{s}, \sigma)$-basis with high enough order σ so $\mathbf{P} = \begin{bmatrix}\mathbf{N}, \bar{\mathbf{N}}\end{bmatrix}$ contains a complete null space basis $\mathbf{N}$ of $\mathbf{F}$. Let k be the column dimension of $\bar{\mathbf{N}}$. Note that the sum of the $\vec{s}$-column degrees of $\mathbf{P}$ is at most $\xi + k\sigma$, since the sum is ξ at order 0 and increases by at most k for each order increase. Also note that the sum of the $\vec{s}$-degrees of $\bar{\mathbf{N}}$ is greater than or equal to the sum of the column degrees of $\mathbf{F}\bar{\mathbf{N}}$, which is at least $k\sigma$ since every column of $\mathbf{F}\bar{\mathbf{N}}$ is nonzero and has order σ. So the sum of the $\vec{s}$-column degrees of $\mathbf{N}$ is bounded by $\xi + k\sigma - k\sigma = \xi$. □

Lemma 2. *Let $\mathbf{F} \in \mathbb{K}[x]^{n/3\times n}$, $\vec{s}$ be a shift with entries bound the corresponding column degrees of $\mathbf{F}$, and $s = \sum \vec{s}/n$ be the average of the entries of $\vec{s}$. Then an $(\mathbf{F}, 4s, \vec{s})$-basis contains at least $2n/5$ null space basis elements.*

Proof. From Lemma 1, the sum of $\vec{s}$-column degrees of any $\vec{s}$-minimal nullspace basis of $\mathbf{F}$ is bounded by $\xi = \sum \vec{s}$. Also, such a nullspace basis has at least $2n/3$ columns so the average $\vec{s}$-column degree is bounded by $3s/2$. Therefore, at least $3/5$ of the columns of such nullspace basis, that is, at least $2n/5$ columns have $\vec{s}$-column degrees bounded by $15s/4$. Since $\mathbf{F}$ multiplied by these columns has the column degrees bounded component-wise by the $\vec{s}$-column degrees of these columns, which are bounded by $15s/4$, an $(\mathbf{F}, 15s/4, \vec{s})$-basis must contain at least $2n/5$ such columns. We can just take the order to be $4s$ instead of $15s/4$ for simplicity. □

Lemma 3. *Let $\vec{s}$, $\mathbf{F}$ be as before. If an $(\mathbf{F}, \sigma, \vec{s})$-basis has columns arranged in increasing $\vec{s}$-column degrees, with $\vec{s}$-column degrees $\vec{b}$. Let $\vec{t} = \vec{b} - \sigma$. Then the $\vec{t}$-column degrees of the basis are bounded component-wise by $\vec{s}$.*

Proof. An $(\mathbf{F}, 0, \vec{s})$-basis has $\vec{s}$-column degrees $\vec{s}$. Then for each order increase, any column of the basis has its $\vec{s}$-column degree increase by at most one, hence at order σ, the $\vec{s}$-column degree increase for each column is at most σ. □

Remark 1. *Note the difference between Lemma 1 and Lemma 3. In Lemma 1, we consider only the shifted degree of a nullspace basis, where as in Lemma 3 applies to a order basis but the order is subtracted from the degrees.*

Lemma 4. *Given $\mathbf{A} \in \mathbb{K}[x]^{m\times n}$, and a shift $\vec{s}$ with entries bound the column degrees of $\mathbf{A}$. Let $\xi = \sum \vec{s}$. Let $\mathbf{B} \in \mathbb{K}[x]^{n\times k}$ and the sum of $\vec{s}$-column degrees of $\mathbf{B}$ is bounded by ξ. Then we can multiply $\mathbf{A}$ and $\mathbf{B}$ with a cost of $O^{\sim}(n^{\omega-1}\xi)$*

Proof. First consider multiplying $\mathbf{A}$ with the columns $\mathbf{B}^{(1)}$ of $\mathbf{B}$ whose $\vec{s}$-degrees are in the range $(\xi/2, \xi]$. There is at most one such column. We do this in $\log n$ steps. At step j we work with the columns $\mathbf{A}_j$ of $\mathbf{A}$ whose column degrees are in the range $(\xi/2^j, \xi/2^{j-1}]$, the the corresponding rows $\mathbf{B}_j^{(1)}$ in $\mathbf{B}^{(1)}$. The column dimension of $\mathbf{A}_j$, which is the same as the row dimension of $\mathbf{B}_j^{(1)}$, is less then 2^j. The degree of $\mathbf{B}_j^{(1)}$ is less than ξ. To use fast multiplication, we expand $\mathbf{B}_j^{(1)}$ to a matrix $\bar{\mathbf{B}}_j^{(1)}$ with degree $\delta \in \Theta(\xi/2^j)$ and column dimension $q \in O(2^j)$ as follows. Write

$$\mathbf{B}_j^{(1)} = \mathbf{B}_{j,0}^{(1)} + \mathbf{B}_{j,1}^{(1)} x^{\delta} + \cdots + \mathbf{B}_{j,q-1}^{(1)} x^{\delta(q-1)} = \sum_{k=0}^{q-1} \mathbf{B}_{j,k}^{(1)} x^{\delta k}$$

with each $\mathbf{B}_{j,k}^{(1)}$ having degree less than δ. Set $\bar{\mathbf{B}}_j^{(1)} = \left[\mathbf{B}_{j,0}^{(1)}, \mathbf{B}_{j,1}^{(1)}, \ldots, \mathbf{B}_{j,q-1}^{(1)}\right]$. We can then multiply $\mathbf{A}_j$, which has dimension $m \times O(2^j)$, and $\bar{\mathbf{B}}_j^{(1)}$, which has dimension $O(2^j)\times O(2^j)$, with a cost of $O^{\sim}\left((m/2^j)\left(2^j\right)^{\omega}\xi/2^j\right) = O^{\sim}\left(\left(2^j\right)^{\omega-2} m\xi\right)$.

Adding up the columns of $\mathbf{A}_j\bar{\mathbf{B}}_j^{(1)}$ gives $\mathbf{A}_j\mathbf{B}_j^{(1)}$ and cost $O(2^j\xi)$. Therefore, multiplying $\mathbf{A}$ and $\mathbf{B}^{(1)}$ over $\log(n)$ steps cost $O^\sim(n^{\omega-2}m\xi)$.

Next we multiply $\mathbf{A}$ with the columns $\mathbf{B}^{(2)}$ of $\mathbf{B}$ whose $\vec{s}$-column degrees are in the range $(\xi/4, \xi/2]$. We proceed in the same way as before, but notice that the columns of $\mathbf{A}$ whose degrees are higher than $\xi/2$ are no longer needed since the corresponding rows in $\mathbf{N}^{(2)}$ must be 0, as the $\vec{t}$-column degree of $\mathbf{N}^{(2)}$ is bounded by $\xi/2$. Multiplying $\mathbf{A}$ and $\mathbf{B}^{(2)}$ also cost $O^\sim(n^{\omega-2}m\xi)$.

Continue doing this, it cost $O^\sim(n^{\omega-2}m\xi)$ to multiply $\mathbf{A}$ with the columns $\mathbf{B}^{(i)}$ of $\mathbf{B}$ whose $\vec{s}$-column degrees are in the range $(\xi/2^i, \xi/2^{i-1}]$. As before, notice that the columns of $\mathbf{A}$ whose degrees are higher than $\xi/2^{i-1}$ are not needed, as $\vec{s}$-column degrees of $\mathbf{B}^{(i)}$ is bounded by $\xi/2^{i-1}$. The overall cost for i from 1 to $\log n$ is $O^\sim(n^{\omega-2}m\xi)$ to multiply $\mathbf{A}$ and $\mathbf{B}$. □

Corollary 1. *The multiplication of* $\mathbf{F}$ *and* $\mathbf{P}_2$ *can be done with a cost of* $O^\sim(n^{\omega-1}\xi)$.

Proof. Since the $\vec{s}$-column degrees $\vec{b}$ of $\mathbf{P}_2$ satisfies $\sum \vec{b} \leq 4sn + \xi = 5\xi$, one way to use Lemma 4 is to use the shift $\vec{s}' = \vec{s} + 4s$. □

Corollary 2. *The multiplication of* $\mathbf{G}_2$ *and* $\mathbf{N}_1$ *can be done with a cost of* $O^\sim(n^{\omega-1}\xi)$.

Proof. Lemma 4 applies directly here. □

Corollary 3. *The multiplication of* $\mathbf{N}_1$ *and* $\mathbf{N}_2$ *can be done with a cost of* $O^\sim(n^{\omega-1}\xi)$.

Proof. Lemma 4 applies because the sum of the column degrees of $\mathbf{N}_1$ is bounded by the sum of the $\vec{t}$-column degrees of $\mathbf{N}_1$, which is $\sum \vec{u} \leq \xi$, and the sum of $\vec{u}$-column degrees of $\mathbf{N}_2$ is also bounded by ξ. □

Lemma 5. *For an input matrix* $\mathbf{F} \in \mathbb{K}[x]^{m\times n}$*, an order vector* $\vec{\sigma}$*, and a shift vector* $\vec{s}$*, if* $\mathbf{P}$ *is a* $(\mathbf{F}, \vec{\sigma}, \vec{s})$*-basis with* $\vec{s}$*-column degrees* $\vec{t}$*, and* $\mathbf{Q}$ *is a* $(\mathbf{FP}, \vec{\tau}, \vec{t})$*-basis with* $\vec{t}$*-column degrees* $\vec{u}$*, where* $\vec{\tau} \geq \vec{\sigma}$ *component-wise, then* $\mathbf{PQ}$ *is a* $(\mathbf{F}, \vec{\tau}, \vec{s})$*-basis with* $\vec{s}$*-column degrees* $\vec{t}$*.*

Proof. It is clear that $\mathbf{PQ}$ has order $(\mathbf{F}, \vec{\tau})$. We now show that $\mathbf{PQ}$ is $\vec{s}$-column reduced and has $\vec{s}$-column degrees $\vec{u}$, or equivalently, $x^{\vec{s}}\mathbf{PQ}$ is column reduced and has column degrees $\vec{u}$. Notice that $x^{\vec{s}}\mathbf{P}$ has column degrees $\vec{t}$ and a full rank leading column coefficient matrix P. Hence $x^{\vec{s}}\mathbf{P}x^{-\vec{t}}$ has column degrees $[0, \ldots 0]$. Similarly, $x^{\vec{t}}\mathbf{Q}x^{-\vec{u}}$ has column degrees $[0, \ldots, 0]$ and a full rank leading column coefficient matrix Q. Therefore, $x^{\vec{s}}\mathbf{P}x^{-\vec{t}}x^{\vec{t}}\mathbf{Q}x^{-\vec{u}} = x^{\vec{s}}\mathbf{PQ}x^{-\vec{u}}$ has column degrees $[0, \ldots, 0]$ and a full rank leading column coefficient matrix PQ. It follows that $x^{\vec{s}}\mathbf{PQ}$ has column degrees $\vec{u}$, or equivalently, the $\vec{s}$-column degrees of $\mathbf{PQ}$ is $\vec{u}$.

It remains to show that any $\mathbf{t} \in \langle(\mathbf{F}, \vec{\tau})\rangle$ is generated by the columns of $\mathbf{PQ}$. Since $\mathbf{t} \in \langle(\mathbf{F}, \vec{\sigma})\rangle$, it is generated by the $(\mathbf{F}, \vec{\sigma})$-basis $\mathbf{P}$, that is, $\mathbf{t} = \mathbf{Pa}$ for $\mathbf{a} = \mathbf{P}^{-1}\mathbf{t} \in \mathbb{K}[x]^n$. Also, $\mathbf{t} \in \langle(\mathbf{F}, \vec{\tau})\rangle$ implies that $\mathbf{a} \in \langle(\mathbf{FP}, \vec{\tau})\rangle$ since

$\mathbf{FPa} = \mathbf{Ft} \equiv 0 \mod x^{\vec{r}}$. It follows that $\mathbf{a} = \mathbf{Qb}$ for $\mathbf{b} = \mathbf{Q}^{-1}\mathbf{a} \in \mathbb{K}[x]^n$. Therefore, $\mathbf{a} = \mathbf{P}^{-1}\mathbf{t} = \mathbf{Qb}$, which gives $\mathbf{t} = \mathbf{PQb}$. □

Lemma 6. *For an input matrix* $\mathbf{G} = \left[\mathbf{G}_1^T, \mathbf{G}_2^T\right]^T \in \mathbb{K}[x]^{m \times n}$ *and a shift vector* $\vec{s}$*, if* $\mathbf{N}_1$ *is a* $\vec{s}$*-minimal nullspace basis of* $\mathbf{G}_1$ *with* $\vec{s}$*-column degrees* $\vec{t}$*, and* $\mathbf{N}_2$ *is a* $\vec{t}$*-minimal nullspace basis of* $\mathbf{G}_2\mathbf{N}_1$ *with* $\vec{t}$*-column degrees* $\vec{u}$*, then* $\mathbf{N}_1\mathbf{N}_2$ *is a* $\vec{s}$*-minimal nullspace basis of* $\mathbf{G}$ *with* $\vec{s}$*-column degrees* $\vec{u}$*.*

Proof. It is clear that $\mathbf{N}_1\mathbf{N}_2$ is a nullspace basis. Showing its $\vec{s}$-column reducedness is same as before. It remains to show that any $\mathbf{n}$ satisfies $\mathbf{Gn} = 0$ must be a linear combination of the columns of $\mathbf{N}_1\mathbf{N}_2$. We use the similar proof as before. First, $\mathbf{n} = \mathbf{N}_1\mathbf{a}$ for some polynomial vector $\mathbf{a}$. Also, $\mathbf{Gn} = 0$ implies that $\mathbf{G}_2\mathbf{N}_1\mathbf{a} = 0$, hence $\mathbf{a} = \mathbf{N}_2\mathbf{b}$ for some vector $\mathbf{b}$. We now have $\mathbf{n} = \mathbf{N}_1\mathbf{N}_2\mathbf{b}$ as required. □