LOW GRADE MATRICES AND MATRIX FRACTION REPRESENTATIONS

Andrew P. Mullhaupt* and Kurt S. Riedel† *

February 18, 2005

Abstract

The lgrade of a $n \times n$ matrix $A$ is the largest rank of any subdiagonal block of a symmetric partition of $A$. A number of algebraic results on lgrade are given. When $A$ has lgrade $d$, it can be approximately decomposed as $A = U + V$, where $U$ is an upper triangular matrix and $V$ has rank $d$. If $A$ satisfies $GA = N$ with $G$ and $N$ have lower bandwidths $d_G$ and $d_N$, then the decomposition is exact: $A = U + V$, where $U$ is an upper triangular matrix with lower bandwidth equal to $d_N - d_G$ and $V$ has low rank (generically $d_G$). This result generalizes well known representations of $A$ when $A = G^{-1}$ and $G$ is banded. A generalization of the Givens rotation product decomposition of unitary Hessenberg matrices is given and its structure analyzed. These ‘consecutive subblock products” are used to construct a representations of a lgrade-d matrix $A$ of the form: $GA = N$ with $G$ and $N$ having lower bandwidth $d$. $G$ can be chosen to be lower triangular or unitary.

AMS Classifications: 15A23

Key words. Matrix Representations, Matrix grade, Matrix Fraction, Matrix Product

---

*S.A.C. Capital Management, LLC, 540 Madison Ave, New York, NY 10022, † Present Address: Millennium Partners, 666 Fifth Ave, New York 10103
1 Introduction.

Let $M$ be a $n \times n$ matrix. We define the lower grade of $M$ as the maximum rank of the lower subdiagonal block of a symmetric partition of $M$. Many common types of special matrices have small lower grade and we prove useful algebraic properties that carry over into large classes of matrices.

In this article, we consider properties of matrices with small lower grade $= d$ and focus on the case $d << n$. We show that lower grade $= d$ matrices may be approximately decomposed $M = U + V$, where $V$ is rank $d$ and $U$ is upper triangular. A $d$-grade matrix, $M$ has two matrix fraction representations: $M = G^{-1}N$ and $M = Q^{-1}N_2$, where $Q$ is unitary and $G$, $N$ and $N_2$ are banded matrices. In deriving these banded matrix fraction representations, we introduce a special set of matrices with low grade: consecutive subblock products. These matrices generalize products of Givens rotations.

We also examine a pseudoconverse of our results by considering the equation $GM = N$, where $G$ and $N$ have bandwidth restrictions. When $N$ is the identity matrix, $M = G^{-1}$, our results correspond to the standard case of band matrices and Green matrices. The special structure of inverses of band matrices is well understood and clearly exposited in [2, 3, 5, 9, 10]. We show that $GM = N$ implies $M$ is the sum of a matrix of specified lower bandwidth and a matrix of specified rank. Our result, Theorem 6.1, requires only that $G$ and $N$ have small lower bandwidth and very mild auxiliary conditions. By combining this result with our matrix fraction representations for grade-$d$ matrices, we show that the decomposition $M = U + V$, where $V$ is rank $d$ and $U$ is upper triangular holds generically for grade-$d$ matrices, where the meaning of ‘generic” is presented in Section 6.

We apply our representation results to an example in signal processing. We now give our basic definitions of bandwidth and grade.

**Definition 1.1** An $n \times n$ matrix $M$ is called lower banded with lower bandwidth (lwidth) $d$ if $M_{ij} = 0$ for $i > j + d$. $M$ is said to have strict lower bandwidth $d$ if $M_{j+d,j} \neq 0$ for $1 \leq j \leq n - d$.

A matrix $T$ is upper triangular if and only if lwidth $T \leq 0$. A matrix with upper bandwidth $1$ is called lower Hessenberg matrix, and a matrix with lower bandwidth $1$ is called an upper Hessenberg matrix.
Definition 1.2 A partition of a matrix

\[ M = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} = \begin{pmatrix} X & Y \\ K & W \end{pmatrix} \]  

(1.1)

is called symmetric if \( X \) and \( W \) are square.

If a symmetric partition exists for \( M \), then \( M \) is square. In this article, we will reserve the symbols \( X, Y, K \) and \( W \) to denote the subblocks of a symmetric partition of \( M \). Our definition of symmetric partition matches that of Rózsa, Romani, Bevilacqua [11]. Their focus is on persymmetric partitions while we derive results for the grade corresponding to symmetric partitions.

Definition 1.3 The upper (lower) grade of a matrix \( M \), written ugrade \((M)\) (lgrade \((M)\)) is the maximum rank of a part of a symmetric partition above (below) the diagonal. The grade of a matrix \( M \) is the maximum rank of an off diagonal part of a symmetric partition, that is, \( \text{grade}(M) = \max\{\text{lgrade}(M), \text{ugrade}(M)\} \).

Proposition 1.1 Companion matrices and Jordan matrices have grade 1. Elementary row (column) operation matrices, Householder, and hyperbolic Householder transformations, Givens and signed Givens rotations all have grade 1.

In the next section, we give properties of matrix grade. Section 3 shows that any matrix \( M \) of lgrade = \( d \) may be approximately decomposed \( M = U + V \) where \( V \) is rank \( d \) and \( U \) is upper triangular. Section 4 describes a special class of low grade matrices: consecutive subblock products. Section 5 shows that any lgrade = \( d \) matrix \( M \) has matrix fraction representation: \( M = G^{-1}N \), where \( G \) and \( N \) are banded matrices.

Section 6 gives a decomposition \( A = U + V \), where \( V \) is low rank and \( U \) has prescribed lwidth under the hypotheses that \( MA = N \) and that \( M \) and \( N \) have prescribed lwidths. In Section 7, we apply our representation results to a class of matrix pairs that are used in signal processing [6, 7].

Notation: The \( n \times n \) identity matrix is \( \mathbb{I}_n \) and the unit vector in the \( k \)th coordinate is denoted \( e_k \). The direct sum of matrices is denoted by \( \oplus \). By \( A_{i:j,k:m} \), we denote the \((j - i + 1) \times (m - k + 1)\) subblock of \( A \) from row \( i \) to row \( j \) and from column \( k \) to column \( m \). We abbreviate \( A_{i:j,1:m} \) by \( A_{i:j} \). We say \( \bar{A} \) is a diagonal subminor of \( A \) if \( \bar{A} = A_{k:m,k:m} \) for some \( k \leq m \).
2 Matrix grade.

We establish some properties of the grade and lgrade. Analogous results hold for upgrade. Some of the proofs are in the Appendix. These algebraic properties are useful in determining or bounding the lgrade of a matrix.

Theorem 2.1 (Basic Properties) Let $M$ be an $n \times n$ matrix:

i) $\text{grade}(M) \leq \lfloor \frac{n^2}{2} \rfloor$.

ii) $\text{grade}(M) \leq \text{rank}(M)$.

iii) $\text{lgrade}(M) \leq \text{lwidth}(M)$, and $\text{grade}(M) \leq \text{width}(M)$.

iv) $\text{lgrade}(M^*) = \text{upgrade}(M)$ and $\text{grade}(M^*) = \text{grade}(M)$.

v) $\text{upgrade}(L) = 0$ is the same as $L$ lower triangular, and $\text{lgrade}(U) = 0$ is the same as $U$ upper triangular. $\text{grade}(M) = 0$ is the same as $M$ diagonal.

vi) Let $N$ be a diagonal subminor of $M$, then $\text{lgrade}(N) \leq \text{lgrade}(M)$. Similarly for grade.

vii) $\text{lgrade}(M_1 \oplus M_2) = \max \{ \text{lgrade}(M_1), \text{lgrade}(M_2) \}$ for square $M_1$ and $M_2$.

This definition of matrix grade is subadditive and submultiplicative:

Theorem 2.2

\[
\text{lgrade}(M_1 + M_2) \leq \text{lgrade}(M_1) + \text{lgrade}(M_2) \tag{2.1}
\]

\[
\text{lgrade}(M_1 M_2) \leq \text{lgrade}(M_1) + \text{lgrade}(M_2). \tag{2.2}
\]

Similarly for upgrade and grade.

The following ‘reduction lemma”, lets us find a succession of matrices which preserve lgrade while transforming $M$. This lemma will be used in the proofs in Section 5.

Lemma 2.3 Let $M$ and $G$ be $n \times n$ matrices with $\text{lgrade}(M) = d$. If for all symmetric partitions of $M$ there exists a matrix $C_i$ such that the corresponding partition of $GM$ satisfies $(GM)_{21} = C_i M_{21}$, then $\text{lgrade}(GM) \leq d$. Only partitions with column dimension $(M_{21}) > d$ and row dimension$(M_{21}) > d$ need be considered.

Proof: Clearly $(GM)_{21} = C_i M_{21}$ has rank less than $d$ so $\text{lgrade}(GM) \leq d$. ■

Typically, $G$ is chosen to be a ‘consecutive subblock product” as defined in Section 4. We now show that the set of matrices of lgrade = $d$ is closed.
Lemma 2.4 \textit{Let }$\|M_k - M\|_F \to 0$\textit{ as }$k \to \infty$, \textit{and }$\lim_{k \to \infty} \operatorname{lgrade}(M_k) = d$. \textit{Then }$\operatorname{lgrade}(M) \leq d$. \textit{Similarly for }$\operatorname{ugrade}$ \textit{and }$\operatorname{grade}$.

Choose any symmetric partition:

$$ M_k = \begin{pmatrix} X_k & Y_k \\ K_k & W_k \end{pmatrix} $$ (2.3)

then $\lim_{k \to \infty} \|K - K_k\|_F = 0$ and $\operatorname{rank}(K) \leq d$ by continuity of the singular values of $K$ in the Frobenius norm. Similarly for $\operatorname{rank}(Y_k)$.

Since any other matrix norm is equivalent to the Frobenius norm, this result holds for all matrix norms. As shown in [11], the grade of $M^{-1}$ equals the grade of $M$:

**Theorem 2.5** [11] \textit{For any square invertible matrix }$M$,

$$ \begin{align*}
\operatorname{lgrade}(M^{-1}) &= \operatorname{lgrade}(M), \\
\operatorname{grade}(M^{-1}) &= \operatorname{grade}(M).
\end{align*} $$ (2.4) (2.5)

**Proof:** Choose any symmetric partition of $M$ as in (1.1) and conformably partition

$$ M^{-1} = \begin{pmatrix} P & Q \\ R & S \end{pmatrix}, $$ (2.6)

so $RX + SK = 0$. If $X$ is nonsingular, then $\operatorname{rank}(R) \leq \operatorname{rank}(K) \leq \operatorname{lgrade}(M)$. This gives $\operatorname{lgrade}(M^{-1}) \leq \operatorname{lgrade}(M)$, but we may equally apply the argument to $M^{-1}$, thus $\operatorname{lgrade}(M^{-1}) = \operatorname{lgrade}(M)$. If $X$ is singular, we obtain the result by adding a small multiple of the identity to $M$ and taking a limit. The result for $\operatorname{ugrade}$ follows by applying this result to $M^*$ and this establishes the result for grade.

**Corollary 2.6** $\operatorname{lgrade}(M^{-1}) \leq \operatorname{lwidth}(M)$, \textit{and }$\operatorname{grade}(M^{-1}) \leq \operatorname{width}(M)$.

In particular, bidiagonal matrices and their inverses have grade 1.

A consequence of submultiplicativity is

**Lemma 2.7 (triangular factor grade)** \textit{Let }$M = NU$ \textit{for upper triangular }$U$. \textit{Then }$\operatorname{lgrade}(M) \leq \operatorname{lgrade}(N)$, \textit{and if either }$U$ \textit{or }$M$ \textit{has nonsingular principal minors then }$\operatorname{lgrade}(M) = \operatorname{lgrade}(N)$.

Note that $M$ nonsingular implies that $U$ is nonsingular and has nonsingular principal minors.
\textbf{Corollary 2.8} Let $M = LR$ for lower triangular $L$ and upper triangular $R$, then $l\text{grade}(M) \leq l\text{grade}(L)$ and $u\text{grade}(M) \leq u\text{grade}(R)$. Additionally, if either $L$ or $R$ is nonsingular, $l\text{grade}(M) = l\text{grade}(R)$ and $u\text{grade}(M) = u\text{grade}(R)$.

Proof: Apply Lemma 2.7 to $M$ and $M^*$. ■

3 Approximate Decomposition for Small Grade Matrices.

We now show that if $l\text{grade}(M) = d$, then for any $\varepsilon > 0$ there is an upper triangular $U_\varepsilon$ and $V_\varepsilon$ of rank $d$ such that $\|M - (U_\varepsilon + V_\varepsilon)\| < \varepsilon$. Here $\|\cdot\|$ is the Frobenius norm. This representation need not be uniform in $\varepsilon$ since $\|U_\varepsilon\|$ and $\|V_\varepsilon\|$ can diverge as $\varepsilon \to 0$. In Sect. 6, we show that this decomposition is exact for ‘generic” grade-$d$ matrices.

We prove this approximation result by constructing a sequence of matrix approximation using the following lemma. The lemma allows us to paste together two overlapping low rank matrices while only altering the upper corner.

\textbf{Lemma 3.1 (The pasting lemma)} Suppose $P$ is a $(j + 1) \times (k + 1)$ matrix:

\[ P = \begin{pmatrix} x^* & y \\ K & w \end{pmatrix}, \tag{3.1} \]

where $K$ is the $j \times k$ lower left submatrix and $\text{rank} \begin{pmatrix} x^* \\ K \end{pmatrix} = d$ and $\text{rank} \begin{pmatrix} K \\ w \end{pmatrix} \leq d$. Then for any $\varepsilon > 0$, there are scalars $\gamma$ and $\delta$ such that

\[ \text{rank} \begin{pmatrix} x^* \\ K + \gamma wx^* \\ \delta \end{pmatrix} \leq d \tag{3.2} \]

and $|\gamma| \|wx^*\|_F < \varepsilon$.

We defer the proof of the pasting lemma until the Appendix. The approximate decomposition is

\textbf{Theorem 3.2} Suppose $M$ is $n \times n$ and $l\text{grade}(M) = d$. Then for any $\varepsilon > 0$ there is an upper triangular $U_\varepsilon$ and $V_\varepsilon$ of rank $d$ such that $\|M - (U_\varepsilon + V_\varepsilon)\| < \varepsilon$. 

6
The idea of the proof is that we repeatedly apply the pasting lemma to make increasingly large subblocks of rank-$d$. At each step the matrix is slightly modified as described in the pasting Lemma. We begin by pasting lower left subblocks whose upper righthand element is along the diagonal. We then do the same trick on each successive superdiagonal. At each step, we show that the process can continue. The formal proof is index rich and we defer it to the Appendix.

In Section 5, we present two matrix fraction representations of matrices with lgrade equal to $d$. The next section introduces the basic building block of this representation.

4 Consecutive subblock products.

As mentioned earlier, many of the basic building blocks for matrix analysis such as elementary row and column operation matrices, signed Givens rotations, and Householder transformations have grade 1. We now present another class of low grade matrices, which we call consecutive subblock products (CSP). In Section 5, we use CSPs to construct representations of grade-$d$ matrices.

**Definition 4.1** Let $F_k$ be an $n \times n$ matrix such that $F_ke_j = e_j$ and $e_j^*F_k = e_j^*$ for $j < k$ and for $j > k + d$. Then $M = F_1 \cdots F_{n-d}$ is called a consecutive subblock product of order $d$.

$F_k$ is the identity except for a $(d+1) \times (d+1)$ block on the diagonal in rows (and columns) $k$ through $k + d$. This block can be chosen arbitrarily. If the $F_k$ are unitary (orthogonal), then so is the product, similarly if the $F_k$ are all lower (upper) triangular, then so is the product.

For the $d = 2$ case our results are summarized by

**Proposition 4.1** ([1]) A product of $n - 1$ consecutive Givens rotations is a unitary Hessenberg matrix, and by the LQ decomposition, a unitary Hessenberg matrix is a product of $n - 1$ consecutive Givens rotations.

Proposition 4.1 follows from ii) and iii) of Theorem 4.2. This representation is extensively studied in [1].

**Theorem 4.2** (Consecutive Subblock Product Properties) Let $M = F_1 \cdots F_{n-d}$ be a consecutive subblock product of order $d$. 

i) If each $F_k$ is nonsingular, then $M^{-*} = (F_1)^{-*} \cdots (F_{n-d})^{-*}$.

ii) Iwidth $M \leq d$. If for each $k$ the element $(F_k)_{k+d,k}$ is nonzero then the consecutive subblock product has strict lower width $d$.

iii) Ugrade $M \leq d$ and grade $M \leq d$.

Proof: ii) is proven by induction. The proof of iii) is in the appendix. The strictness result in ii) of Theorem 4.2 is useful in analyzing Corollary 6.2.

## 5 Band fraction representations.

In this section, we show that if lgrade $M \leq d$, then there are matrices $G$ and $H$ such that $GM = H$, where $G$ and $H$ have band structure. If $M$ satisfies a generic condition, $G$ is invertible and we have a structured matrix fraction representation: $M = G^{-1}H$. Similarly we have a second matrix fraction representation $M = Q^{-1}H_2$, where $Q$ is unitary. Both $G$ and $Q$ are consecutive subblock products with a very special structure.

**Theorem 5.1** Suppose lgrade($M$) $\leq d$. Then there are matrices $L$ and $H$ such that $LM = H$ with $L$ lower triangular and lwidth $L \leq d$ and lwidth $H \leq d$ and uwidth $H \leq$ uwidth $M$.

The proofs in this section are inductive and begin by partitioning $M$ as (1.2), where $M_{22}$ is $(d+1) \times (d+1)$. Since lgrade $M \leq d$, then rank $M_{21} \leq d$. A matrix is given which eliminates the last row of $M_{21}$ while preserving the lgrade of the product.

Proof: The result is trivially true for $n \leq d$. Suppose it is true for $n-1$. Partition $M$ as in (1.2). Since lgrade $M \leq d$, then rank $M_{21} \leq d$, there exists a nonzero $(d+1)$ vector $\alpha$ such that $\alpha^*M_{21} = 0$. Let $\hat{F}$ be any lower triangular $(d+1) \times (d+1)$ matrix with last row $\alpha^*$, for example $\hat{F} = I_d + e_d(\alpha - e_d)^*$. We define $F_{n-d-1} = I_{n-d-1} \oplus \hat{F}$. The last row of $\hat{F}M_{21}$ vanishes. By the Lemma 2.3, the rank of the first $d$ rows of $\hat{F}M_{21}$ have rank less than or equal to $d$. This reduces the problem to the claimed factorization of the $(n-1) \times (n-1)$ leading minor of $M$, from which one determines $F_{n-d-2}$, ..., $F_1$, then put $L = F_1 \cdots F_{n-d-1}$. ■

**Corollary 5.2** In Theorem 5.1, $L$ may be chosen to be a consecutive subblock product with $F_i = I_n + e_{i+d}(\alpha_i - e_{i+d})^*$, where $e_{i+d}$ is unit vector in the $(i+d)$ th coordinate.
Suppose in the last proof, that all the requisite vectors \( \alpha_i \) can be taken to have a nonzero last element, then \( L \) is nonsingular and \( M = L^{-1}H \). The following condition guarantees the existence of such \( \alpha_i \):

**Definition 5.1 (Condition \( \Gamma \))** Let \( M \) be a \( n \times n \) matrix and define the \((d + 1) \times (d + 1)\) submatrices \( M_i \) by \( M_i \equiv M_{(i;i + d),1:(i + 1)} \) for \( 1 \leq i < n - d \). The matrix \( M \) satisfies condition \( \Gamma \) if and only if for all \( i \) with \( 1 \leq i < n - d \), \( e_{d+1} \notin \text{Range}(M_i) \), where the \((d + 1)\)-vector \( e_{d+1} \equiv (0, 0 \ldots 1)^* \).

Condition \( \Gamma \) implies that for each \( i \) there exists a nonnull \((d + 1)\)-vector \( \alpha_i \) such that \( \alpha_i^* M_i = 0 \) and \( \alpha_{i,d+1} \neq 0 \) (by the duality \( \mathcal{N}(M_i^*)^\perp = \mathcal{R}(M_i) \)).

**Corollary 5.3** If \( \lgrade M \leq d \) and \( M \) satisfies Condition \( \Gamma \), then \( L \) is nonsingular and \( M = L^{-1}H \) where \( L \) and \( H \) are given in Theorem 5.1. \( F_k \) and \( L \) may be chosen to be unit lower triangular.

**Theorem 5.4** Suppose \( \lgrade M \leq d \). There are matrices \( Q \) and \( H \) such that \( QM = H \) where \( \lwidth Q \leq d \), \( \lwidth H \leq d \) and \( Q \) is unitary consecutive subblock product with each subblock, \( F_k \) being a Householder transformation. If \( M \) is real, then \( \text{real orthogonal} \).

Proof: Repeat the proof of Theorem 5.1 except this time choose \( F_{n-d-1} \) to be the Householder reflection determined \( \hat{F}_{n-d-1} \alpha = e_{d+1} \).

**Theorem 5.5** Suppose \( \lgrade M \leq d \). Then there are matrices \( Q \) and \( H \) such that \( MQ = H \) with \( Q \) is an unitary consecutive subblock product of width \( d \) (and \( \text{real orthogonal if } M \text{ is real} \) and \( \lwidth Q \leq d \) and \( \lwidth H \leq d \).

Proof: The result is trivially true for \( n \leq d \). Suppose it is true for \( n - 1 \). There is a \((d + 1) \times (d + 1)\) unitary matrix \( F_1 \) such that the first column of \( M_{21}F_1 \) vanishes. (For example, there exists a nonzero vector \( \alpha \) such that \( M_{21} \alpha = 0 \), let \( F_1 \) be the Householder reflection such that \( F_1 \alpha = e_1 \).) Trivially, the rank of the first \( d \) rows of \( M_{21}F \) have rank less than or equal to \( d \), and this reduces the problem to the \((n - 1) \times (n - 1)\) submatrix of \( M \) below the first row and column, which determines \( F_2, \ldots, F_{n-d-1} \). Put \( Q = F_1 \cdots F_{n-d-1} \). Transposing the theorems and proofs gives similar results for upper graded matrices. In the constructions of Theorems 5.4 and 5.5, \( F_k \) can be a product of \( d \) Givens rotations instead of a Householder reflection.
6 \quad M = U + V \text{ decomposition with low lwidth } U \text{ and low rank } V.

The inverse of a strict invertible lower bidiagonal matrix, $M$, is the sum of a matrix with $\text{lwidth} = -1$ and a rank-1 matrix [2, 3, 5, 9, 10]. In our notation, these authors study $M$, where $GM = I$ and $G$ has low lwidth. We now study the generalized problem: $GM = N$. We show that the decomposition $M = U + V$, where $V$ is low rank and $U$ has small lwidth holds when $G$ and $N$ have small lwidth and other mild assumptions.

**Theorem 6.1** Suppose $\text{lwidth}(G) = d_G$ and $\text{lwidth}(N) = d_N$, and

$$GM = N. \quad (6.1)$$

Then $M = U + V$ where $\text{lwidth} U = d_N - d_G$ and $\text{rank}(V) \leq \text{dim ker} \, T$, where $T$ is the upper triangular matrix and $Z$ is the lower shift matrix and

$$T = \begin{cases} Z^{d_G} G & d_G \geq 0 \\ Z^{-d_G} G & d_G < 0 \end{cases}. \quad (6.2)$$

When $G$ is strict, then $\text{rank}(V) = d_G$.

**Proof:** Suppose $d_G \geq 0$. Then

$$Z^{d_G} GM = TM = Z^{d_G} N \quad (6.3)$$

for the upper triangular matrix $T = Z^{d_G} G$. Let $R = T + W$ be nonsingular upper triangular with rank $W = \text{dim ker} \, T$, then

$$M = R^{-1} TM + R^{-1} WM = R^{-1} Z^{d_G} N + R^{-1} WM. \quad (6.4)$$

Define $U = R^{-1} Z^{d_G} N$ and $V = R^{-1} WM$. Since $\text{lwidth} R^{-1} Z^{d_G} N \leq d_N - d_G$, we have $M = U + V$ where $\text{lwidth} U \leq d_N - d_G$ and $\text{rank} \, V \leq \text{dim ker} \, T$.

If $d_G < 0$, modify this proof by replacing $Z^{d_G}$ with $Z^{-d_G}$ everywhere in the last paragraph. $\blacksquare$

**Definition 6.1** An $n \times n$ matrix $M$ has an additive decomposition of order $d$ when $M = U + V$, where $U$ is upper triangular and $\text{rank}(V) \leq d$. 

Clearly, if $M$ has an additive decomposition of order $d$, then it has grade-$d$. Theorem 3.2 shows that the set of matrices with additive decomposition of order $d$ is dense in the set matrices of grade-$d$.

Combining Theorem 6.1 with Theorem 5.4 yields

**Corollary 6.2** Suppose $\mathrm{lgrade} M \leq d$. Then there are unitary $Q$ with $\mathrm{lwidth}(Q) = d$ and $H$ with $\mathrm{lwidth}(H) = d$, such that $QM = H$ and $M = U + V$, where $U$ is upper triangular, $\mathrm{rank}(V) = \dim \ker(Z^d Q)$. If $Q$ has strict $\mathrm{lwidth} d$, then $\mathrm{rank}(V) \leq d$.

**Proof:** Apply Theorem 6.1 to the factorization $QM = H$ given by Theorem 5.4. □

Theorem 6.1 does not use matrix grade, but Corollary 6.2 does. We now show that a ‘generic” grade-$d$ matrix satisfies this additive decomposition. Our analysis is on a set of matrix pairs which define grade-$d$ matrices.

**Definition 6.2** Let $\Omega_d$ be the set of $n \times n$ matrix pairs $(G, H)$, where $G$ and $H$ have $\mathrm{lwidth}-d$ and $G$ is invertible. Let $\Omega_0^d$ be the subset of $\Omega_d$ where $G$ is strict.

Theorem 5.4 implies that every grade-$d$ matrix $M$ is defined by at least one matrix pair $(G, H)$ in $\Omega_d$. Theorem 6.1 states that every matrix pair $(G, H)$ in $\Omega_0^d$ defines an unique matrix $M$ such that $M = G^{-1} H$ and $M$ has an additive decomposition, $M = U + V$, of order $d$. Thus every element in $\Omega_0^d$ corresponds to a grade-$d$ matrix.

Since the set of nonstrict matrix pairs in $\Omega_d$ has codimension one in $\Omega_d$, we interpret our results as meaning that a generic grade-$d$ matrix has an additive decomposition. We note this is generic on the covering set, $\Omega_d$. Since many matrix pairs in $\Omega_0^d$ could generate the same grade-$d$ matrix, we cannot make definitive statements on the genericity of the additive decomposition on the set of grade-$d$ matrices itself.

If we so desire, we can restrict both $\Omega_d$ and $\Omega_0^d$ by the requirement that $G$ be an unitary matrix which is a consecutive subblock product of Householder rotations. This and other restrictions will help to reduce the multiplicity of matrix fraction representation in $\Omega_d$ that correspond to the same grade-$d$ matrix.

We skip the hard analysis of determining a subset of matrix fraction representations that is in exact one to one correspondence with the set of grade-$d$ matrices. Corollary 6.2 says that if the $Q$ used in the construction of Theorem 5.4 is strict, then there is an additive decomposition of order $d$. We interpret ‘genericity” as meaning that the set of strict $Q$ is generic in the set of all such $Q$. 

11
7 Triangular input normal matrices.

This section applies the previous results to a special class of matrix pairs used in signal processing [6, 7]. The concatenation of such a matrix pair is an orthogonal matrix. Thus we begin by deriving results on the grade of an unitary matrix.

Theorem 7.1 If \( M \) is unitary then \( \text{lgrade}(M) = \text{ugrade}(M) \).

Proof: Since \( M^{-1} = M^* \), then \( \text{lgrade}(M) = \text{lgrade}(M^{-1}) = \text{ugrade}(M) \).

More generally, Theorem 7.1 applies to hyperexchange matrices. The \( d = 1 \) case yields an approximate converse of Proposition 4.1:

Corollary 7.2 Let \( M \) be unitary and Hessenberg, then \( \text{grade}(M) = 1 \).

As described in [6], triangular input normal (TIN) pair are useful for system identification. In [7], some TIN pairs are analyzed and a matrix fraction representation is given. We now show that TIN pairs have low-grade.

Definition 7.1 A matrix pair is \( (A, B) \) is a lower triangular input normal pair if and only if

\[
AA^* + BB^* = I, \tag{7.1}
\]

where \( A \) is a \( n \times n \) lower triangular, \( B \) is size \( n \times d \), and \( I \) is the identity matrix.

Theorem 7.3 Let \( (A, B) \) be a lower TIN system with \( B \) of full column rank. Then \( \text{grade}(A) \leq \text{rank}(B) \).

Proof: By Theorem 2.5.1 of [4], there are \( K \) and \( W \) such that \( M = \begin{pmatrix} A & B \\ K & W \end{pmatrix} \) is \( (n + d) \times (n + d) \) and \( M \) is unitary. By Theorem 2.1, part vi) and Theorem 7.1, \( \text{grade}(A) \leq \text{grade}(M) = \text{ugrade}(M) \). Choose any symmetric partition of \( M \) and let \( M_{12} = M_{1:k,(k+1):((n+d)} \). If \( k \geq n \), then \( \text{rank}(M_{12}) \leq d \). If \( k < n \), then \( M_{12} = (0_{1:k,(k+1):n} | B_{1:k,:}) \), where \( 0_{1:k,(k+1):n} \) is a matrix of all zeros. Thus \( \text{rank}(M_{12}) \leq d \).

In [7], we give, an explicit matrix fraction representation of TIN matrix pairs with \( d = 1 \), \( A = M^{-1}N \), where \( M \) and \( N \) are bidiagonal lower triangular matrices. Theorem 6.1 implies that \( A \) is the sum of a diagonal matrix and the lower half of a rank one matrix.

We now apply Theorem 5.4 to TIN pairs.
Corollary 7.4 Suppose $(A, B)$ be a TIN pair of grade $d$. There are matrices $Q$ and a lower triangular $H$ such that $M = QA$ and $B = Q\tilde{B}$, where $Q$ is unitary consecutive subblock product of grade $d$, $\tilde{B}_{(d+1):n,n} = 0$ and $MM^* = I_n - \tilde{B}\tilde{B}$.

Proof: We follow the notation in Theorem 5.4. In the subblock decomposition of $Q$, we choose $F_{n-d-1}$ to be the Householder reflection that zeros out $B_{n,:}$. We then choose $F_{n-d-2}$ to zero out the $(n-1)$st row of $F_{n-d-1}B$. We continue this procedure until $QB_{(d+1):n,n} = 0$. Conjugating (7.1) by $Q$ completes the result. ■

Similarly, Theorem 5.1 may be applied to TIN pairs to yield another matrix fraction representation [7].

8 Summary

The lgrade of a matrix satisfies a number of useful algebraic properties. Our results are particularly useful when the lgrade of a matrix is significantly smaller than the matrix dimension. Section 5 gives two matrix fraction decompositions of low lgrade matrices: $M = L^{-1}H$ and $M = Q^{-1}H$, where $L$ and $H$ have smal lower bandwidth and $Q$ is unitary. Both $L$ and $Q$ are consecutive subblock products. We have shown that state space representations can be constructed from these matrix fraction representations.

Any matrix $M$ with lgrade $(M) = d$ may be approximated to arbitrary accuracy as the sum: $M = U + V$, where $U$ is upper triangular $U$ and $V$ is rank $d$. When $GM = N$ and with small lower bandwidth $G$ and $N$, the decomposition $M = U + V$ is exact where the lwidth of $U$ and the rank of $V$ are given in Theorem 6.1. This result generalize the classic analysis of the inverse of a banded matrix. The hypotheses of Theorem 6.1 are mild and do not use the matrix grade. By combining Theorem 6.1 with the matrix fraction decompositions, we show that generically a grade-$d$ matrix has an exact additive decomposition of order $d$: $M = U + V$ exactly.

9 Appendix: Proofs

Proof of Theorem 2.1:
Part iii) Choose any symmetric partition as in (1.1). $K$ is upper triangular, with at most width $(M)$ nonzero columns, so rank $(K) \leq \text{width}(M)$, similarly for $Y$. 

13
Part vi) Let $K$ be a subdiagonal part of a symmetric partition of $N$ of maximal rank. Since $N$ is obtained by deleting rows and columns of $M$, either $K$, or $K$ with columns and rows adjoined, is a subdiagonal part of a symmetric partition of $M$. Hence $\operatorname{lgrade}(N) \leq \operatorname{rank}(K) \leq \operatorname{lgrade}(M)$.

Proof of Part vii): Let

$$M_1 \oplus M_2 = \begin{pmatrix} M_1 & 0 \\ 0 & M_2 \end{pmatrix} = \begin{pmatrix} X & Y \\ K & W \end{pmatrix}$$

be any symmetric partition of $M_1 \oplus M_2$. Either $K = \begin{pmatrix} K_1 \\ 0 \end{pmatrix}$ where $K_1$ is a subdiagonal block of $M_1$ or $K = \begin{pmatrix} 0 & K_2 \end{pmatrix}$ where $K_2$ is a subdiagonal block of $M_2$. Hence

$$\operatorname{lgrade}(M_1 \oplus M_2) \leq \max \{ \operatorname{lgrade}(M_1), \operatorname{lgrade}(M_2) \},$$

(9.2)

but both $M_1$ and $M_2$ are minors of $M_1 \oplus M_2$ so

$$\operatorname{lgrade}(M_1 \oplus M_2) \geq \max \{ \operatorname{lgrade}(M_1), \operatorname{lgrade}(M_2) \}.$$

(9.3)

Proof of Theorem 2.2: Choose any symmetric partition

$$M_i = \begin{pmatrix} X_i & Y_i \\ K_i & W_i \end{pmatrix}.$$ 

(9.4)

Then

$$M_1 + M_2 = \begin{pmatrix} X_1 + X_2 & Y_1 + Y_2 \\ K_1 + K_2 & W_1 + W_2 \end{pmatrix}$$

(9.5)

and $\operatorname{rank}(K_1 + K_2) \leq \operatorname{lgrade}(M_1) + \operatorname{lgrade}(M_2)$.

Since $\operatorname{rank}(K_i) \leq d_i$ and

$$M_1M_2 = \begin{pmatrix} X_1X_2 + Y_1K_2 & X_1Y_2 + Y_1W_2 \\ K_1X_2 + W_1K_2 & K_1Y_2 + W_1W_2 \end{pmatrix},$$

(9.6)

we have $\operatorname{rank}(K_1X_2 + W_1K_2) \leq \operatorname{lgrade}(M_1) + \operatorname{lgrade}(M_2)$, hence $\operatorname{lgrade}(M_1M_2) \leq \operatorname{lgrade}(M_1) + \operatorname{lgrade}(M_2)$. ■

Proof of Lemma 2.7: $\operatorname{lgrade}(M) \leq \operatorname{lgrade}(N) + \operatorname{lgrade}(U) = \operatorname{lgrade}(N)$. Choose a symmetric partition

$$N = \begin{pmatrix} X_N & Y_N \\ K_N & W_N \end{pmatrix} \quad \text{and} \quad U = \begin{pmatrix} X_U & Y_U \\ 0 & W_U \end{pmatrix},$$

(9.7)
so that $K_N$ has maximal rank, that is, $\text{rank}(K_N) = \text{lgrade}(N)$. Then the corresponding partition of $M$ is

$$
M = \begin{pmatrix}
X_NX_U & X_NY_U + Y_NW_U \\
K_NX_U & K_NY_U + W_NW_U
\end{pmatrix}.
$$

(9.8)

If $U$ has nonsingular leading minors, then $X_U$ is nonsingular and so $\text{lgrade}(N) = \text{rank}(K_N) \leq \text{lgrade}(M)$. Now if $M$ has nonsingular leading minors then $X_NX_U$ is nonsingular, hence $X_U$ is nonsingular.

Proof of Lemma 3.1: Suppose that $w$ is in the column space of $K$, then $w = Ku$ for some $u$. (If $w \equiv 0$, then $u \equiv 0$.) Set $\gamma = 0$ and $\delta = x^*u$, and verify

$$
\begin{pmatrix}
x^* & \delta \\
K & w
\end{pmatrix} = \begin{pmatrix} x^* \\
K \end{pmatrix} \begin{pmatrix} u \\
I \end{pmatrix}.
$$

(9.9)

Alternatively, suppose $w$ is not in the span of the columns of $K$. Then rank $K < d$, but rank $\begin{pmatrix} x^* \\
K + \gamma wx^* \end{pmatrix} = d$. $K$ has at least $d$ columns, and let $u^*$ be the project of $x^*$ perpendicular to the span of the rows of $K$. Note that $w \not\equiv 0$ and $u \not\equiv 0$. Put $\gamma = (2\|wx^*\|_F)^{-1} \varepsilon$, $h = 1/(\gamma x^*u)$, $\delta = hx^*u = 2\|wx^*\|_F/\varepsilon$, and verify

$$
\begin{pmatrix}
x^* & \delta \\
K + \gamma wx^* & w
\end{pmatrix} = \begin{pmatrix} x^* \\
K + \gamma wx^* \end{pmatrix} \begin{pmatrix} u \\
hu \end{pmatrix}.
$$

(9.10)

The left matrix factor has a rank of at most $d$, so the left hand side has a rank at most $d$.

Proof of Theorem 3.2: Let $\|M\|_L \equiv \sum_{i>j} |M_{ij}|^2$. We define a sequence of matrices, $\{M^{(k)}|0 \leq k \leq f\}$ such that $M^{(0)} = M$, $M^{(f)} = V_\varepsilon$ and $\|M^{(k+1)} - M^{(k)}\|_L < \epsilon/f$. At each step, $M^{(k+1)} - M^{(k)}$ is given by applying the pasting lemma to a lower left subblock of $M^{(k)}$. At each step, let $P^{(k)}$ be the $(n-i+1) \times j$ subblock of $M^{(k-1)}$: $P^{(k)} = M^{(k-1)}_{(i,n,1;j)}$, where $i(k)$ and $j(k)$ are indices which we will specify shortly.

We begin by applying the pasting lemma to each subblock with upper right corner on the main diagonal. This implies $i(k) = k + 1$ and $j(k) = k + 1$ for $1 \leq k < n - 1$. Next we apply the pasting lemma to each subblock with upper right corner on the first superdiagonal. $i(k) = k - n - 2$ and $j(k) = i(k) + 1$ for $n - 1 \leq k < 2n - 2$. The pasting lemma is then applied to each successive superdiagonal.

The proof is by finite induction on the superdiagonals of $M^{(k)}$. The lower left subblock of $P^{(k)}$, $M^{(k-1)}_{(i+1,n,1;j-1)}$ has rank of at most $d$. Let $\tilde{P}^{(k)}$ be the result of applying
the pasting lemma to $P^{(k)}$. Set $M^{(k)}_{i:n,1:j} = \tilde{P}^{(k)}$ and otherwise $M^{(k)} = M^{(k-1)}$. The modified matrix subblock, $\tilde{P}^{(k)}_{(i+1):n,1:j}$ spans the same column space as $P^{(k)}_{(i+1):n,1:j}$. The same result holds for $\tilde{P}^{(k)}_{i:n,1:(j-1)}$. Thus the pasting does not increase the rank of any lower left subblock of $M^{(k-1)}$ that arises from a symmetric partition or that has been processed in an earlier step. Therefore we can continue the process until we reach the upper right corner of the matrix. The final matrix, $M^{(f)}$ has rank $d$ or less and differs from the original matrix by an upper triangular matrix and a modification of size $\epsilon$. ■

Proof of iii) of Theorem 4.2:

By induction. It is clearly true for $n = d + 1$. Suppose then, that the result is true for $(n-1) \times (n-1)$ consecutive subblock products of order $d$. Now choose any symmetric partition of $M$ as in (1.2). We decompose $M$ as the product of $F_1$ with $F_2F_3\ldots F_{n-d}$ and denote the analogous symmetric partition of $F_2F_3\ldots F_{n-d}$ by $\{X_2, Y_2, K_2, W_2\}$:

$$
M = \left(\Phi \oplus I \ 0 \right) \left( X_2 \ Y_2 \right) = \left(\Phi \oplus I \ 0 \right) \left( 1 \ 0 \right) \left( 0 \ M_2 \right).
$$

(9.11)

where $\Phi$ is the $(d + 1) \times (d + 1)$ principal subminor of $F_1$, and $\Phi \oplus I$ conforms to $X_2$. Here $M_2$ is an $(n-1) \times (n-1)$ consecutive subblock product of order $d$, we have grade $M_2 \leq d$. Now grade $(1 \oplus M_2) \leq d$, so rank $Y_2 \leq d$

$$
\text{rank } Y = \text{rank } (\Phi \oplus I) Y_2 \leq \text{rank } Y_2 \leq d.
$$

(9.12)

The result follows from the reduction lemma. ■

References


