LAA Project

Afrad Basheer, Akhil Naravi, Nikhil Vishwanath

April 10, 2019

Contents

1 Introduction

Matrix factorizations or decompositions are fundamentally important in providing practical numerical algorithms and theoretical linear algebra insights. Matrix factorizations are examples of perhaps the most important strategy of numerical analysis: replace a relatively difficult problem with a much easier one, e.g. triangular systems are easier to solve than full systems.

The paper "*A RANK-ONE REDUCTION FORMULA AND ITS APPLICATIONS TO MATRIX FACTORIZATIONS*" by *Chu, Funderlic and Golub* shows that it is possible to unify matrix factorizations with a common algorithm by exploiting a remarkably simple and powerful idea of rank reduction due to Wedderburn. We have organised the main ideas from this paper and simplified several theorems and proofs to better understand these ideas.

2 Wedderburn's Fundamental Idea

2.1 Rank Reduction Theorem

Wedderburn observed that rank one matrices of the form $\omega^{-1} \cdot (Ax)(y^t A)$ when subtracted from A result in a matrix of rank one less than A if $\omega = y^t Ax \neq 0$. Successive use of this idea will provide general factorizations of A. We first state and prove the fundamental theorem by Wedderburn.

Theorem 2.1.1. *Suppose* $A \in \mathbb{R}^{m \times n},\ x \in \mathbb{R}^n,\ y \in \mathbb{R}^m$ *such that* $\omega = y^t Ax \neq 0$. *Then*

$$\mathit{rank}\Big(A - \omega^{-1} \cdot (Ax)(y^t A)\Big) = \mathit{rank}(A) - 1 \tag{2.1.1}$$

Proof. Let $B = A - \omega^{-1} \cdot (Ax)(y^t A)$. We will prove that $\text{kernel}(B) = \text{kernel}(A) \oplus \text{span}(\{x\})$. Then, by the Rank-Nullity theorem we have,

$$\text{rank}(A) + \text{null}(A) = n = \text{rank}(B) + \text{null}(B) = \text{rank}(B) + \text{null}(A) + 1 \implies \text{rank}(A) = \text{rank}(B) + 1$$

Observe that $\text{kernel}(A) \subseteq \text{kernel}(B) : Az = 0 \implies Bz = Az - \omega^{-1} \cdot (Ax)y^t(Az) = 0$. If $Bz = 0$ then $0 = Az - \omega^{-1} \cdot (Ax)(y^t Az) = A(z - k \cdot x)$ where $k = \omega^{-1}(y^t Az) \in \mathbb{R}$. Now, $(z - k \cdot x) \in \text{kernel}(A) \implies z \in \text{kernel}(A) + \text{span}(\{x\})$. Finally, since z was arbitrary and $Ax \neq 0$, we have $\text{kernel}(B) = \text{kernel}(A) \oplus \text{span}(\{x\})$ □

2.2 Factorization

Having realized that equation 2.1.1 is a rank reducting operation, we can apply it successively so as to get a sequence of Wedderburn Matrices $A_k \neq 0$ such that $A_k := A_{k-1} - \omega_{k-1}^{-1} \cdot (A_{k-1}x_{k-1})(y_{k-1}^t A_{k-1})$ as long as $\omega_{k-1} = y_{k-1}^t A_{k-1} x_{k-1} \neq 0$. Thus we get the following matrix outer-product factorization:

$$A = \begin{bmatrix} A_1x_1 & A_2x_2 & \dots & A_\gamma x_\gamma \end{bmatrix} \begin{bmatrix} \omega_1^{-1} & 0 & & \dots & 0 \\ 0 & \omega_2^{-1} & 0 & \dots & 0 \\ & 0 & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ 0 & 0 & \dots & 0 & \omega_\gamma^{-1} \end{bmatrix} \begin{bmatrix} y_1^t A_1 \\ y_2^t A_2 \\ \vdots \\ \vdots \\ y_\gamma^t A_\gamma \end{bmatrix} \tag{2.2.1}$$

3 Biconjugation Process

The biconjugation process gives us the algorithm to arrive at a pair of matrices that will be used to factorize matrices. This pair is called the biconjugate pair and is defined below.

Definition 3.0.1. *For any $A \in \mathbb{R}^{m\times n}$, $U \in \mathbb{R}^{n\times k}$, $V \in \mathbb{R}^{m\times k}$, (U, V) is a biconjugate pair (with respect to A) if $\Omega := V^t AU$ is non-singular and diagonal. Such a decomposition is called a biconjugate decomposition.*

3.1 Definitions

The following definitions are used throughout the biconjugation process.

D.1 $\langle x, y\rangle_A := y^t Ax$ for all $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$

D.2 $A_1 := A$ and for $k \in \{2, \ldots, \gamma = \text{rank}(A)\}$, $A_k := A_{k-1} - \omega_{k-1}^{-1} \cdot (A_{k-1}x_{k-1})(y_{k-1}^t A_{k-1})$ where $x_{k-1} \in \mathbb{R}^n$ and $y_{k-1} \in \mathbb{R}^m$ such that $\omega_{k-1} = y_{k-1}^t A_{k-1}x_{k-1} \neq 0$

D.3 $u_1 := x_1$ and $v_1 := y_1$ so that $\langle u_1, v_1\rangle_A \neq 0$

D.4 $u_2 := x_2 - \left(\frac{\langle x_2,v_1\rangle_A}{\langle u_1,v_1\rangle_A}\right) \cdot u_1$ and $v_2 := y_2 - \left(\frac{\langle u_1,y_2\rangle_A}{\langle u_1,v_1\rangle_A}\right) \cdot v_1$ (well defined by **D.3**)

D.5 For all $x \in \{x_1, \ldots, x_\gamma\}$ and for all $y \in \{y_1, \ldots, y_\gamma\}$,

$$\mathcal{S}(x, y) := \left\{ z - \left(\frac{\langle z, y\rangle_A}{\langle x, y\rangle_A}\right) \cdot x \;\middle|\; z \in \mathbb{R}^n \right\} \quad \text{and} \quad \mathcal{T}(x, y) := \left\{ w - \left(\frac{\langle x, w\rangle_A}{\langle x, y\rangle_A}\right) \cdot y \;\middle|\; w \in \mathbb{R}^m \right\}$$

3.2 Facts

The following theorems, which are encountered in basic courses in linear algebra, will be used in the proofs that follow but will not be proved here.

F.1 $\text{kernel}(A) = \text{range}(A^t)^\perp$

F.2 $\text{range}(A) = \text{kernel}(A^t)^\perp$

F.3 $W \subseteq V \implies V^\perp \subseteq W^\perp$

3.3 Observations

We make the following observations using the definitions and facts stated above:

O.1 $x_k \in \text{kernel}(A_{k+1}) = \text{range}(A_{k+1}^t)^\perp$

Proof. $A_{k+1}x_k = A_kx_k - \omega^{-1} \cdot (A_kx_k)(y_k^t A_kx_k) = A_kx_k - \omega^{-1} \cdot (A_kx_k)\omega = A_kx_k - A_kx_k = 0 \implies x_k \in \text{range}(A_{k+1}^t)^\perp$ by **F.1** □

O.2 $y_k \in \text{kernel}(A_{k+1}^t) = \text{range}(A_{k+1})^\perp$

Proof. $y_k^t A_{k+1} = y_k^t A_k - \omega^{-1} \cdot (y_k^t A_kx_k)(y_k^t A_k) = y_k^t A_k - \omega^{-1}\omega \cdot (y_k^t A_k) = y_k^t A_k - y_k^t A_k = 0 \implies y_k \in \text{range}(A_{k+1})^\perp$ by **F.2** □

O.3 $\text{kernel}(A_2) = \text{kernel}(A_1) \oplus \text{span}(\{x_1\})$ (by proof of Theorem 2.1.1)

O.4 For all $x \in \{x_1, \ldots, x_\gamma\}$ and for all $y \in \{y_1, \ldots, y_\gamma\}$,

$$\mathbb{R}^n = \mathrm{span}(\{x\}) \oplus \mathcal{S}(x,y) \quad \text{and} \quad \mathbb{R}^m = \mathrm{span}(\{y\}) \oplus \mathcal{T}(x,y)$$

Proof.

- $\forall\, z \in \mathbb{R}^n,\ z = \left(z - \left(\frac{\langle z,y\rangle_A}{\langle x,y\rangle_A}\right)\cdot x\right) + \left(\frac{\langle z,y\rangle_A}{\langle x,y\rangle_A}\right)\cdot x = z_1 + z_2$ where $z_1 \in \mathcal{S}(x,y)$ and $z_2 \in \mathrm{span}(\{x\})$ which implies that $\mathbb{R}^n = \mathrm{span}(\{x\}) + \mathcal{S}(x,y)$
- If $z \in \mathrm{span}(\{x\}) \cap \mathcal{S}(x,y)$ then

$$z = \alpha \cdot x = z_1 - \left(\frac{\langle z_1, y\rangle_A}{\langle x, y\rangle_A}\right)\cdot x \text{ for some } \alpha \in \mathbb{R} \text{ and } z_1 \in \mathbb{R}^n \tag{3.3.1}$$

Multiplying by A_2 we get $0 = \alpha \cdot A_2 x = A_2 z_1 - \left(\frac{\langle z_1,y\rangle_A}{\langle x,y\rangle_A}\right)\cdot A_2 x = A_2 z_1 - 0$ (by **O.1**) and hence $z_1 \in \mathrm{kernel}(A_2)$, which by **O.3** means that either $z_1 \in \mathrm{kernel}(A_1)$ or $z_1 = k \cdot x$.

If $z_1 \in \mathrm{kernel}(A_1)$ then $\langle z_1, y\rangle_A = 0$ and multiplying 3.3.1 by A_1 gives $\alpha \cdot A_1 x = A_1 z_1 + \left(\frac{\langle z,y\rangle_A}{\langle x,y\rangle_A}\right)\cdot A_1 x = 0$ and since $A_1 x \neq 0$ this means that $\alpha = 0 \implies z = 0$

If $z_1 = k \cdot x$ then $z_1 - \left(\frac{\langle z_1,y\rangle_A}{\langle x,y\rangle_A}\right)\cdot x = k\cdot x - \left(\frac{\langle x,y\rangle_A}{\langle x,y\rangle_A}\right) k \cdot x = 0 \implies z = 0$

Hence $\mathrm{span}(\{x\}) \cap \mathcal{S}(x,y) = \{0\}$

□

O.5 $\forall\, z \in \mathbb{R}^n,\ A_{k+1} z = A_k \left(z - \left(\frac{\langle z, y_k\rangle_A}{\langle x_k, y_k\rangle_A}\right)\cdot x_k\right) = A_k w$ for some $w \in \mathcal{S}(x_k, y_k)$

Proof. $A_{k+1} z = A_k z - \omega^{-1}\cdot(A_k x_k)(y_k^t A_k z) = A_k z - (y_k^t A_k x_k)^{-1}(y_k^t A_k z)(A_k x_k) = A_k z - \left(\frac{\langle z,y_k\rangle_A}{\langle x_k,y_k\rangle_A}\right)\cdot A_k x_k = A_k \left(z - \left(\frac{\langle z,y_k\rangle_A}{\langle x_k,y_k\rangle_A}\right)\cdot x_k\right)$ □

O.6 $\mathrm{range}(A_{k+1}) \subsetneq \mathrm{range}(A_k)$ (follows from **O.5** and the fact that $\mathcal{S}(x_k, y_k) \subsetneq \mathbb{R}^n$)

O.7 $A u_2 = A_2 x_2$ and $v_2^t A = y_2^t A_2$ (follows from **O.5**)

O.8 $\langle u_2, v_1\rangle_A = \langle u_1, v_2\rangle_A = 0$

Proof. $\langle u_2, v_1\rangle_A = v_1^t(A u_2) = v_1^t(A_2 x_2)$ (by **O.7**) $= 0$ (by **O.2**). Similarly, $\langle u_1, v_2\rangle_A = (v_2^t A) u_1 = (y_2^t A_2) u_1$ (by **O.7**) $= 0$ (by **O.1**) □

O.9 $\omega_2 = y_2^t A_2 x_2 = \langle u_2, v_2\rangle_A$

Proof. $y_2^t A_2 x_2 = v_2^t A x_2$ (by **O.7**) $= v_2^t A \left(u_2 + \left(\frac{\langle x_2,v_1\rangle_A}{\langle u_1,v_1\rangle_A}\right)\cdot u_1\right) = v_2^t A u_2 + \left(\frac{\langle x_2,v_1\rangle_A}{\langle u_1,v_1\rangle_A}\right)\cdot v_2^t A u_1 = v_2^t A u_2$ since $v_2^t A u_1 = 0$ (by **O.8**) □

3.4 Biconjugation Algorithm

This algorithm constructs a biconjugate pair (U, V) using the matrix pair (X, Y), where $X = \left[x_1 \mid \cdots \mid x_\gamma\right]$ and $Y = \left[y_1 \mid \cdots \mid y_\gamma\right]$.

Theorem 3.4.1. *Suppose $A \in \mathbb{R}^{m\times n}$ has rank γ. Let $\{x_1, \ldots, x_\gamma\}$ and $\{y_1, \ldots, y_\gamma\}$ be any vectors associated with the Webberburn Rank-Reducing process (i.e. $y_k^t A_k x_k \neq 0$ for each k). Then for all $k \in \{2, \ldots, \gamma\}$,*

$\mathfrak{T}$*1* $u_k := x_k - \sum_{i=1}^{k-1} \left(\frac{\langle x_k, v_i\rangle_A}{\langle u_i, v_i\rangle_A}\right) \cdot u_i$ *and* $v_k := y_k - \sum_{i=1}^{k-1} \left(\frac{\langle u_i, y_k\rangle_A}{\langle u_i, v_i\rangle_A}\right) \cdot v_i$ *are well defined*

$\mathfrak{T}$*2* $Au_k = A_k x_k$ *and* $v_k^t A = y_k^t A_k$

$\mathfrak{T}$*3* $\langle u_k, v_j\rangle_A = \langle u_j, v_k\rangle_A = 0$ *for all* $j < k$

$\mathfrak{T}$*4* $\omega_k = y_k^t A_k x_k = \langle u_k, v_k\rangle_A$

Proof.
When $k = 2$, **$\mathfrak{T}$1**, **$\mathfrak{T}$2**, **$\mathfrak{T}$3**, **$\mathfrak{T}$4** are equivalent to **D.4**, **O.7**, **O.8**, **O.9** respectively. Now suppose the theorem is true for $j \leq k < \gamma$.

Then $\omega_k = \langle u_k, v_k\rangle_A \neq 0$, so **$\mathfrak{T}$1** is true for $k + 1$ and

$$\begin{aligned}
A_{k+1} &= A_k - \omega_k^{-1} \cdot (A_k x_k)(y_k^t A_k) && \text{(by D.2)}\\
&= A_k - \omega_k^{-1} \cdot (Au_k)(v_k^t A) && \text{(by } \mathfrak{T}2\text{)}\\
&= A - \sum_{i=1}^{k} \omega_i^{-1} \cdot (Au_i)(v_i^t A) && \text{(by recursion)}
\end{aligned}$$

Which gives us

$$\begin{aligned}
Au_{k+1} &= A\left(x_{k+1} - \sum_{i=1}^{k}\left(\frac{\langle x_{k+1}, v_i\rangle_A}{\langle u_i, v_i\rangle_A}\right)\cdot u_i\right) && \text{(by } \mathfrak{T}1 \text{ for } k+1\text{)}\\
&= Ax_{k+1} - \sum_{i=1}^{k}\left(\frac{\langle x_{k+1}, v_i\rangle_A}{\langle u_i, v_i\rangle_A}\right)\cdot Au_i\\
&= Ax_{k+1} - \sum_{i=1}^{k} \omega_i^{-1}\cdot (Au_i)(v_i^t A x_{k+1}) && \text{(by definitions of } \langle x_{k+1}, v_i\rangle_A \text{ and } \langle u_i, v_i\rangle_A\text{)}\\
&= \left(A - \sum_{i=1}^{k}\omega_i^{-1}\cdot (Au_i)(v_i^t A)\right) x_{k+1}\\
&= A_{k+1} x_{k+1}
\end{aligned}$$

and,

$$
\begin{aligned}
v_{k+1}^t A &= \left(y_{k+1}^t - \sum_{i=1}^{k} \left(\frac{\langle u_i, y_{k+1} \rangle_A}{\langle u_i, v_i \rangle_A} \right) \cdot v_i^t \right) A && \text{(by } \mathfrak{T}\mathbf{1} \text{ for } k+1) \\
&= y_{k+1}^t A - \sum_{i=1}^{k} \left(\frac{\langle u_i, y_{k+1} \rangle_A}{\langle u_i, v_i \rangle_A} \right) \cdot v_i^t A \\
&= y_{k+1}^t A - \sum_{i=1}^{k} \omega_i^{-1} \cdot (y_{k+1}^t A u_i)(v_i^t A) && \text{(by definitions of } \langle u_i, y_{k+1} \rangle_A \text{ and } \langle u_i, v_i \rangle_A) \\
&= y_{k+1}^t \left(A - \sum_{i=1}^{k} \omega_i^{-1} \cdot (A u_i)(v_i^t A) \right) \\
&= y_{k+1}^t A_{k+1}
\end{aligned}
$$

Hence, $\mathfrak{T}\mathbf{2}$ is true for $k+1$.

By the definitions of u_k and v_k,

$$
\operatorname{span}\{x_1, \ldots, x_j\} = \operatorname{span}\{u_1, \ldots, u_j\} \tag{3.4.1}
$$

$$
\operatorname{span}\{y_1, \ldots, y_j\} = \operatorname{span}\{v_1, \ldots, v_j\} \tag{3.4.2}
$$

Since, $y_j \in \operatorname{range}(A_{j+1})^\perp$ for $1 \le j \le k$ (**O.2**) and $\operatorname{range}(A_{k+1}) \subsetneq \operatorname{range}(A_k) \subsetneq \cdots \subsetneq \operatorname{range}(A_1)$ (**O.6**), it follows that $\operatorname{range}(A_{k+1})^\perp \supsetneq \operatorname{range}(A_k)^\perp \supsetneq \cdots \supsetneq \operatorname{range}(A_1)^\perp$ (by **F.3**) and hence

$$
y_j \in \operatorname{range}(A_{k+1})^\perp \implies v_j \in \operatorname{range}(A_{k+1})^\perp \quad \text{for } 1 \le j < k+1 \quad \text{(by 3.4.2)} \tag{3.4.3}
$$

and similarly $$x_j \in \operatorname{range}(A_{k+1}^t)^\perp \implies u_j \in \operatorname{range}(A_{k+1}^t)^\perp = \operatorname{kernel}(A_{k+1}) \quad \text{for } 1 \le j < k+1 \quad \text{(by 3.4.1)} \tag{3.4.4}$$

Then for $1 \le j < k+1$ we get,

$$
\langle u_{k+1}, v_j \rangle_A = v_j^t (A u_{k+1}) = v_j^t (A_{k+1} x_{k+1}) \text{ (by } \mathfrak{T}\mathbf{2}) = 0 \text{ (by 3.4.3)} \tag{3.4.5}
$$

$$
\langle u_j, v_{k+1} \rangle_A = (v_{k+1}^t A) u_j = (y_{k+1}^t A_{k+1}) u_j \text{ (by } \mathfrak{T}\mathbf{2}) = v_{k+1}^t (A_{k+1} u_j) = 0 \text{ (by 3.4.4)} \tag{3.4.6}
$$

Hence, $\mathfrak{T}\mathbf{3}$ is true for $k+1$.

Finally,

$$
\begin{aligned}
\omega_{k+1} &= y_{k+1}^t A_{k+1} x_{k+1} && (\mathbf{D.2}) \\
&= v_{k+1}^t A \left(u_{k+1} + \sum_{i=1}^{k} \left(\frac{\langle x_{k+1}, v_i \rangle_A}{\langle u_i, v_i \rangle_A} \right) \cdot u_i \right) && \text{(by } \mathfrak{T}\mathbf{2} \text{ and } \mathfrak{T}\mathbf{1}) \\
&= v_{k+1}^t A u_{k+1} + \sum_{i=1}^{k} \left(\frac{\langle x_{k+1}, v_i \rangle_A}{\langle u_i, v_i \rangle_A} \right) \cdot v_{k+1}^t A u_i \\
&= v_{k+1}^t A u_{k+1} + 0 && \text{(by 3.4.6)} \\
&= \langle u_{k+1}, v_{k+1} \rangle_A
\end{aligned}
$$

Hence, $\mathfrak{T}\mathbf{4}$ is true for $k+1$. $\square$

3.5 Decomposition and Factorization

The following corollary uses the biconjugation algorithm to define biconjugate decomposition, biconjugate factorization and gives the relationship between (X, Y) and (U, V).

Corollary 3.5.1. *Suppose $A \in \mathbb{R}^{m\times n}$ has rank γ. Let $\{x_1, \ldots, x_\gamma\}$ and $\{y_1, \ldots, y_\gamma\}$ be any vectors associated with the Wedderburn Rank-Reducing process (i.e. $y_k^t A_k x_k \neq 0$ for each k).*

For $k \in \{1, \ldots, \gamma\}$, define

- $\Omega_k :=$ *diag*$(\omega_1, \ldots, \omega_k)$ *where* $\omega_k := y_k^t A_k x_k$
- $U_k := [u_1 \mid \cdots \mid u_k] \in \mathbb{R}^{n\times k}$ *and* $V_k = [v_1 \mid \cdots \mid v_k] \in \mathbb{R}^{m\times k}$ *where* $u_k := x_k - \sum_{i=1}^{k-1} \left(\frac{\langle x_k, v_i \rangle_A}{\langle u_i, v_i \rangle_A} \right) \cdot u_i$ *and* $v_k := y_k - \sum_{i=1}^{k-1} \left(\frac{\langle u_i, y_k \rangle_A}{\langle u_i, v_i \rangle_A} \right) \cdot v_i$

Then,

$\mathfrak{C}$*1* $V_k^t A U_k = \Omega_k$ *for* $1 \leq k \leq \gamma$ *(Biconjugate Decomposition)*

$\mathfrak{C}$*2* $A = AU_\gamma \Omega_\gamma^{-1} V_\gamma^t A$ *(Biconjugate Factorization)*

$\mathfrak{C}$*3* *There exist unique unit upper triangular matrices $R_x, R_y \in \mathbb{R}^{k\times k}$ such that $X_k = U_k R_x$ and $Y_k = V_k R_y$*

Proof.

- **$\mathfrak{C}$1** follows from **O.8** and **O.9**
- **$\mathfrak{C}$2** follows from equation 2.2.1 and **$\mathfrak{T}$2**
- **$\mathfrak{C}$3** follows from **$\mathfrak{T}$1** since the jth column of R_x is

$$\left[\frac{\langle x_j, v_1 \rangle_A}{\langle u_1, v_1 \rangle_A}, \ldots, \frac{\langle x_j, v_{j-1} \rangle_A}{\langle u_{j-1}, v_{j-1} \rangle_A}, 1, 0 \ldots 0 \right]^t$$

  and the jth column of R_y is

$$\left[\frac{\langle u_1, y_j \rangle_A}{\langle u_1, v_1 \rangle_A}, \ldots, \frac{\langle u_{j-1}, y_j \rangle_A}{\langle u_{j-1}, v_{j-1} \rangle_A}, 1, 0 \ldots 0 \right]^t$$

  They are unique because each of $\{x_1, \ldots, x_k\}, \{y_1, \ldots, y_k\}, \{u_1, \ldots, u_k\}$ and $\{v_1, \ldots, v_k\}$ are linearly independent.

□

3.6 Characterizing Biconjugatability I

The biconjugation algorithm is very strongly connected to Wedderburn's Rank Reducing process: we need to get the vectors x_k and y_k such that ${y_k}^t A_k x_k \neq 0$, for which we need the Wedderburn matrices A_k. We will now proceed to characterize biconjugatability in order to decouple it from Wedderburn's matrices. We want to be able to arrive at a biconjugate pair (U, V) without having to generate the Wedderburn matrices.

Definition 3.6.1. *Given $A \in \mathbb{R}^{m\times n}$, a pair of matrices $(X, Y) \in \mathbb{R}^{n\times k} \times \mathbb{R}^{m\times k}$ is said to be biconjugatable, and biconjugated into a biconjugate pair of matrices (U, V), if there exist unit upper triangle matrices $R_x, R_y \in \mathbb{R}^{k\times k}$ such that $X = UR_x$ and $Y = VR_y$.*

Lemma 3.6.1. *If $A \in \mathbb{R}^{n\times n}$ has a unit LDU decomposition, i.e. $A = LDU$ where $L \in \mathbb{R}^{n\times n}$ is a lower-triangular matrix with ones on its diagonal, $U \in \mathbb{R}^{n\times n}$ is a upper-triangular matrix with ones on its diagonal and $D \in \mathbb{R}^{n\times n}$ is diagonal, then the decomposition is unique.*

Proof. Suppose $A = LDU = L'D'U'$. Then $(L'^{-1}L)D = D'(U'U^{-1})$. Now since both L and L' are lower-triangular, L'^{-1} is lower-triangular and hence $L'^{-1}LD$ is also lower-triangular. Similarly, $D'U'U^{-1}$ is upper-triangular. A lower-triangular matrix which equals an upper-triangular matrix is diagonal and vice-versa. So $L'^{-1}LD$ and $D'U'U^{-1}$ are both diagonal, which means that $L'^{-1}L$ and $U'U^{-1}$ are diagonal. Moreover, both $L'^{-1}L$ and $U'U^{-1}$ have ones on the diagonal (because L, L', U, U' have ones on the diagonal) and hence both are the identity matrix. Thus $L = L'$, $U = U'$ and $D = D'$. □

Theorem 3.6.1. *Suppose $A \in \mathbb{R}^{m\times n}$ and $(X, Y) \in \mathbb{R}^{n\times k} \times \mathbb{R}^{m\times k}$. Then (X, Y) can be biconjugated, with respect to A, into a biconjugate pair of matrices (U, V) if and only if Y^tAX has a unit LDU decomposition. In which case,*

- *U, V, Ω of the biconjugate decomposition are unique*
- *The LDU Factorization of Y^tAX is $R_y^t\Omega R_x$ where R_x, R_y are the unique unit upper triangular matrices such that $X = UR_x$ and $Y = VR_y$*

Proof. Suppose (X, Y) is biconjugatable. Then (by definition 3.6.1) there exist unit upper triangular matrices R_x, R_y such that $X = UR_x$, $Y = VR_y$ and $V^tAU = \Omega$ is a non-singular diagonal matrix. It follows that $Y^tAX = (R_y^tV^t)A(UR_x) = R_y^t(V^tAU)R_x = R_y^t\Omega R_x$ is a unit triangular LDU decomposition of Y^tAX, which is unique.

Conversely, if $Y^tAX = R_2^t\Omega R_1$ is a unit LDU decomposition such that both R_1, R_2 are unit upper triangular matrices, then $({R_2^t}^{-1}Y^t)A(XR_1^{-1}) = {R_2^t}^{-1}(Y^tAX)R_1^{-1} = {R_2^t}^{-1}(R_2^t\Omega R_1)R_1^{-1} = ({R_2^t}^{-1}R_2^t)\Omega(R_1R_1^{-1}) = \Omega$ is non-singular and diagonal and hence, (X, Y) biconjugates into $(XR_1^{-1}, YR_2^{-1}) = (U, V)$ (Note that, $X = UR_1$ and $Y = VR_2$). Since the LDU decomposition is unique, R_1, R_2 are unique and hence U, V, Ω are unique. □

4 Factorizations of Matrices using the Biconjugation Process

4.1 Trapezoidal LDU

Lemma 4.1.1. *$B \in \mathbb{R}^{n\times n}$ has a unit LDU factorization if and only if det$(B[1:k,1:k]) \neq 0$ for $1 \leq k \leq n$.*

Theorem 4.1.1. *Suppose $A \in \mathbb{R}^{m\times n}$ has rank γ and det$(A[1:k,1:k]) \neq 0$ for $1 \leq k \leq \gamma$. Then there exist upper-trapezoidal pair of matrices $(X_\gamma Y_\gamma) \in \mathbb{R}^{n\times\gamma} \times \mathbb{R}^{m\times\gamma}$ that are biconjugatable into the biconjugate pair (U_γ, V_γ) such that $AU_\gamma \in \mathbb{R}^{m\times\gamma}$ and $A^tV_\gamma \in \mathbb{R}^{n\times\gamma}$ are lower-trapezoidal matrices and $A = (AU_\gamma){\Omega_\gamma}^{-1}(A^tV_\gamma)^t$ (the biconjugate factorization) is a trapezoidal LDU factorization of A.*

Proof. Let X_γ be the first γ columns of the identity matrix in $\mathbb{R}^{n\times n}$ and Y_γ be the first γ columns of the identity matrix in $\mathbb{R}^{m\times m}$. Then $Y_\gamma^t A X_\gamma = A[1:\gamma, 1:\gamma]$, which by lemma 4.1.1 has a unit LDU factorization. So $(X_\gamma Y_\gamma)$ are biconjugatable. Let (U_γ, V_γ) be the biconjugate pair obtained by the applying the biconjugate algorithm to $(X_\gamma Y_\gamma)$.

From the biconjugation algorithm, we have the following results:

1. $AU_\gamma = \left[A_1x_1 \mid A_2x_2 \mid \cdots \mid A_\gamma x_\gamma\right]$ and $A^tV_\gamma = \left[A_1^t y_1 \mid A_2^t y_2 \mid \cdots \mid A_\gamma^t y_\gamma\right]$.
2. $x_k \in \text{kernel}(A_{k+1}) \subsetneq \text{kernel}(A_{k+2}) \subsetneq \cdots \subsetneq \text{kernel}(A_\gamma)$ (**O.1**) for $k < \gamma$ and hence
$$\{x_1, \ldots, x_{k-1}\} \in \text{kernel}(A_k) \text{ for } 2 \leq k \leq \gamma$$
3. $y_k \in \text{kernel}({A_{k+1}}^t) \subsetneq \text{kernel}({A_{k+2}}^t) \subsetneq \cdots \subsetneq \text{kernel}({A_\gamma}^t)$ (**O.2**) for $k < \gamma$ and hence
$$\{y_1, \ldots, y_{k-1}\} \in \text{kernel}({A_k}^t) \text{ for } 2 \leq k \leq \gamma$$

By our choice of x_ks (columns of X_γ) and y_ks (columns of Y_γ) and results (2) and (3), it follows that the first $k-1$ rows and columns of A_k are 0. So the first $k-1$ elements of the vectors A_kx_k and ${A_k}^ty_k$ are 0 and hence AU_γ and A^tV_γ are lower-trapezoidal matrices.

It is now clear that the biconjugate factorization of A, i.e. $A = (AU_\gamma){\Omega_\gamma}^{-1}(A^tV_\gamma)^t$, is a trapezoidal LDU factorization of A. ☐

4.2 Cholesky Factorization

Theorem 4.2.1. *If $A \in \mathbb{R}^{n\times n}$ is symmetric and positive definite, then the Cholesky Factorization of A can be obtained by applying the biconjugation algorithm to (I_n, I_n)*

Proof. Since ${I_n}^tAI_n = A$ satisfies $\det(A[1:k,1:k]) > 0$ for $1 \leq k \leq n$, ${I_n}^tAI_n$ has an LDU factorization and hence (I_n, I_n) can be biconjugated into a biconjugate pair (U, V).

Because A is symmetric, each Wedderburn matrix A_k is also symmetric: ${A_{k+1}}^t = {A_k}^t - {\omega_k}^{-1}\cdot\big((A_ke_k)({e_k}^tA_k)\big)^t = {A_k}^t - {\omega_k}^{-1}\cdot({A_k}^te_k)({e_k}^t{A_k}^t) = A_k - {\omega_k}^{-1}\cdot(A_ke_k)({e_k}^tA_k) = A_{k+1}$. This means that,

$$\begin{aligned}
AV &= A^tV && (A \text{ is symmetric}) \\
&= \left[A^tv_1 \mid \cdots \mid A^tv_n\right] \\
&= \left[{A_1}^te_1 \mid \cdots \mid {A_n}^te_n\right] && (\text{by } \mathfrak{T}\mathbf{2}) \\
&= \left[A_1e_1 \mid \cdots \mid A_ne_n\right] && (\text{each } A_k \text{ is symmetric}) \\
&= \left[Au_1 \mid \cdots \mid Au_n\right] && (\text{by } \mathfrak{T}\mathbf{2}) \\
&= AU \\
\implies V &= U && (A \text{ is invertible})
\end{aligned}$$

Moreover, by $\mathfrak{C}$**3**, there exist unit upper-triangular matrices R_x, R_y such that $UR_x = I_n$ and $VR_y = I_n$. So both $U = {R_x}^{-1}$ and $V = {R_y}^{-1}$ are unit upper-triangular matrices. The biconjugate decomposition then gives us $\Omega = V^tAU = U^tAU$ and hence, $A = U^{-t}\Omega U^{-1} = (U^{-t}\sqrt{\Omega})(\sqrt{\Omega}U^{-1}) = (U^{-t}\sqrt{\Omega})(U^{-t}\sqrt{\Omega})^t$, where $U^{-t}\sqrt{\Omega}$ is a lower-triangular matrix, is the Cholesky Factorizatin of A. □

4.3 Characterizing Biconjugatability II

Theorem 4.3.1. *The matrix pair (X, Y) is biconjugatabe into the biconjugate pair (U, V) (with respect to some $A \in \mathbb{R}^{m\times n}$) if and only if there exist unique unit upper-triangular matrices R_x and R_y satisfying $X = UR_x$ and $Y = VR_y$.*

Proof. $\mathfrak{C}$**3** gives us the forward implication. Conversely, suppose (U, V) is a biconjugatable pair and R_x, R_y are unit upper-triangular matrices such that $X = UR_x$ and $Y = VR_y$. Then $(VR_y)^tA(UR_x) = {R_y}^t(V^tAU)R_x = {R_y}^t\Omega R_x$ is an LDU factorization of itself. So by theorem 3.6.1, (X, Y) can be biconjugated into (U, V). □

4.4 QR Decomposition

Theorem 4.4.1. *Suppose $A \in \mathbb{R}^{m\times n}$ ($m \geq n$) has full rank (n). Then the reduced QR Factorization of A can be obtained by applying the biconjugation algorithm to (I, A).*

Proof. If $A = QR$ is the reduced QR factorization of A then we may write it as $A = Q\Psi R_1$ where Ψ is a diagonal matrix and R_1 is a *unit* upper triangular matrix. Then $(Q\Psi)^tA{R_1}^{-1} = \Psi^t(Q^tA{R_1}^{-1}) = \Psi^t\Psi = \Psi^2$ is diagonal and non-singular. So by definition 3.0.1, $({R_1}^{-1}, Q\Psi)$ is a biconjugate pair, which we denote as (U, V).

Now by theorem 4.3.1, we know that (X, Y) biconjugates into (U, V) if there exists unit upper-triangular matrices R_x, R_y such that $X = UR_x$ and $Y = VR_y$. Letting $R_x = R_y = R_1$, we get $X = {R_1}^{-1}R_1 = I$ and $Y = Q\Psi R_1 = A$.

Therefore, we can apply the biconjugate algorithm to (I, A) to get (U, V) which gives $V^tAU = \Psi^2$ and hence $Q = V(V^tAU)^{-1/2}$ and $R = (V^tAU)^{-1/2}U^{-1}$. □