Wedderburn's Rank Reduction Theorem

Afrad Basheer
Akhil Naravi
Nikhil Vishwanath

M.Sc. Data Science
Chennai Mathematical Institute

Introduction

"A rank one reduction formula and its applications to matrix factorization"
by Moody T. Chu, Robert. E. Funderlic and Gene H. Golub

The Purpose of this paper is to unify matrix factorizations and several fundamental linear algebra processes within a common framework by exploiting a remarkably simple powerful idea of rank reduction due to Wedderburn.

Wedderburn's Rank Reduction Theorem

Theorem

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 $\text{rank}\left(A - \omega^{-1} \cdot (Ax)(y^t A)\right) = \text{rank}(A) - 1$.

Proof.

Let $B = A - \omega^{-1} \cdot (Ax)(y^t A)$.
We will prove that $\text{kernel}(B) = \text{kernel}(A) \oplus \text{span}(\{x\})$.

By the Rank-Nullity theorem we get,
$\text{rank}(A) = \text{rank}(B) + 1$

Observe that,
$\text{kernel}(A) \subseteq \text{kernel}(B)$:
$Az = 0 \implies Bz = Az - \omega^{-1} \cdot (Ax)y^t(Az) = 0.$

□

Proof.

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 $\text{kernel}(B) = \text{kernel}(A) \oplus \text{span}(\{x\})$ □

Wedderburn Rank Reduction

Rank Reducing process

$A_{k+1} := A_k - {\omega_k}^{-1} A_k x_k {y^T}_k A_k$

Matrices A_k are called Wedderburn matrices. Where $A_1 = A$.
By applying this process repeatedly we can get sequence of Wedderburn matrices.
Which can be summarised as $A = \Phi\Omega^{-1}\Psi^T$. Where,

$$\Omega := diag[\omega_1, \omega_2, ...], \Phi = [\phi_1, \phi_2 ...] \in R^{m\times\gamma} \quad \text{and}$$
$$\Psi := [\psi_1, \psi_2, ..] \in R^{n\times\gamma}$$
$$\text{with,} \quad \phi_k := A_k x_k$$
$$\psi_k := {A_k}^T y_k$$

Wedderburn Rank Reduction

Matrix Outer-Product Factorization

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

Biconjugation Process

Definition

Let $A \in R^{m \times n}$, $U \in R^{n \times k}$ and $V \in R^{m \times k}$. Then (U, V) is a **Biconjugate pair** with respect to A if : $\Omega := V^T AU$

Definition

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

Facts

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

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

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

Biconjugation Process

Definition

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

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

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

3: $\text{kernel}(A_2) = \text{kernel}(A_1) \oplus \text{span}(\{x_1\})$

Biconjugation Process

Definition

For all $x \in \{x_1, \ldots, x_\gamma\}$, $\mathcal{S}(x, y) := \left\{ z - \left(\frac{\langle z, y \rangle_A}{\langle x, y \rangle_A} \right) \cdot x \;\middle|\; z \in \mathbb{R}^n \right\}$

O.4: For all $x \in \{x_1, \ldots, x_\gamma\}$, $\mathbb{R}^n = \text{span}(\{x\}) \oplus \mathcal{S}(x, y)$

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 $w \in \mathcal{S}(x_k, y_k)$

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

Biconjugation Process

Definitions

- $u_1 := x_1$ and $v_1 := y_1$ so that $\langle u_1, v_1 \rangle_A \neq 0$
- $u_2 := x_2 - \left(\frac{\langle x_2, v_1 \rangle_A}{\langle u_1, v_1 \rangle_A} \right) \cdot u_1$ and $v_2 := y_2 - \left(\frac{\langle u_1, y_2 \rangle_A}{\langle u_1, v_1 \rangle_A} \right) \cdot v_1$

O.7: $Au_2 = A_2 x_2$ and $v_2^t A = y_2^t A_2$ (follows from Obs 5)

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

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

Theorem

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 $k \in \{2, \ldots, \gamma\}$,

$\mathfrak{T}$**2.1** $$u_k := x_k - \sum_{i=1}^{k-1} \left(\frac{\langle x_k, v_i \rangle_A}{\langle u_i, v_i \rangle_A} \right) \cdot u_i \text{ and}$$
$$v_k := y_k - \sum_{i=1}^{k-1} \left(\frac{\langle u_i, y_k \rangle_A}{\langle u_i, v_i \rangle_A} \right) \cdot v_i \text{ are well defined}$$

$\mathfrak{T}$**2.2** $Au_k = A_k x_k$ *and* $v_k^t A = y_k^t A_k$

$\mathfrak{T}$**2.3** $\langle u_k, v_j \rangle_A = \langle u_j, v_k \rangle_A = 0$ *for all* $j < k$

$\mathfrak{T}$**2.4** $\omega_k = y_k^t A_k x_k = \langle u_k, v_k \rangle_A$

Biconjugation Algorithm

Proof

- When $k = 2$, $\mathfrak{T}1, \mathfrak{T}2, \mathfrak{T}3, \mathfrak{T}4$ are equivalent to the defintion of u_2, v_2, O7, O8, O9 respectively
- If the theorem is true for $j \leq k < \gamma$,
- $\omega_k = \langle u_k, v_k \rangle_A \neq 0$ so $\mathfrak{T}1$ is true for $k + 1$

  $$A_{k+1} = A_k - \omega_k^{-1} \cdot (Au_k)(v_k^t A) = A - \sum_{i=1}^{k} \omega_i^{-1} \cdot (Au_i)(v_i^t A)$$

- which gives us $Au_{k+1} = A_{k+1}x_{k+1}$ and $v_{k+1}{}^t A = y_{k+1}{}^t A_k$, ($\mathfrak{T}2$)

  by definitions $\mathfrak{T}1$, $\text{span}\{x_1, \ldots, x_j\} = \text{span}\{u_1, \ldots, u_j\}$ and $\text{span}\{y_1, \ldots, y_j\} = \text{span}\{v_1, \ldots, v_j\}$
- $v_j \in \text{range}(A_{k+1})^\perp$ for $1 \leq j < k + 1$ and $u_j \in \text{range}(A_{k+1}^t)^\perp = \text{kernel}(A_{k+1})$ for $1 \leq j < k + 1$
- thus $\langle u_{k+1}, v_j \rangle_A = \langle u_j, v_{k+1} \rangle_A = 0$ for all $j < k + 1$ ($\mathfrak{T}3$) and $\omega_{k+1} = y_{k+1}^t A_{k+1} x_{k+1} = \langle u_{k+1}, v_{k+1} \rangle_A$ ($\mathfrak{T}4$)

Decomposition and Factorisation

Corollary

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

$$\Omega_k := \operatorname{diag}(\omega_1, \ldots, \omega_k) \text{ where } \omega_k := y_k^t A_k x_k$$

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

Then,

- $\mathfrak{C}\mathbf{1}$ $V_k^t A U_k = \Omega_k$ for $1 \leq k \leq \gamma$ (Biconjugate Decomposition)
- $\mathfrak{C}\mathbf{2}$ $A = AU_\gamma \Omega_\gamma^{-1} V_\gamma^t A$ (Biconjugate Factorization)
- $\mathfrak{C}\mathbf{3}$ 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 $Y_k = V_k R_y$

Characterizing Biconjugatibility

Definition

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 $Y = VR_y$.

Theorem

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 $Y^t AX$ has a unit LDU decomposition. In which case,

- U, V, Ω of the biconjugate decomposition are unique
- The LDU Factorization of $Y^t AX$ 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 $Y = VR_y$

Recap

- $A_{k+1} := A_k - {\omega_k}^{-1} A_k x_k {y^T}_k A_k$
- Let $A \in R^{m \times n}$, $U \in R^{n \times k}$ and $V \in R^{m \times k}$. Then (U, V) is a **Biconjugate pair** with respect to A if : $\Omega := V^T AU$

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 $Y^t AX$ has a unit LDU decomposition. In which case,

- U, V, Ω of the biconjugate decomposition are unique
- The LDU Factorization of $Y^t AX$ 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 $Y = VR_y$

Lemma

$B \in \mathbb{R}^{n\times n}$ 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

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 $A^t V_\gamma \in \mathbb{R}^{n\times\gamma}$ are lower-trapezoidal matrices and $A = (AU_\gamma)\Omega_\gamma{}^{-1}(A^t V_\gamma)^t$ (the biconjugate factorization) is a trapezoidal LDU factorization of A.

Theorem

If $A \in \mathbb{R}^{n \times n}$ is symmetric and positive definite and (X_γ, X_γ) is biconjugatable, then the resulting biconjugate pair(U_γ, V_γ) has $U_\gamma = V_\gamma$.

Theorem

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)

Theorem

If $(U, V) \in R^{n\times k} \times R^{m\times k}$ is a biconjugate pair , and R_x and R_y are arbitrary unit upper triangular matrices in $R^{k\times k}$, then (UR_x, VR_y) can be biconjugated and the resulting biconjugate pair is exactly the initial (U, V).

Theorem

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).

Theorem

Let $\sigma_1 \geq \sigma_2 \geq .... \geq \sigma_r$ be singular values of A. Then $||A_k||_2 = \sigma_k$

Theorem

Let x_k, y_k be singular vectors of A_k, then they can also be singular vectors of A.