Wedderburn Decomposition and Its Applications to Matrix Factorizations

Moody T. Chu
Department of Mathematics
North Carolina State University
Raleigh, North Carolina 27695-8205

Joint work with R. E. Funderlic

Outline

- Rank Reduction Formula
- Wedderburn Process
- Inverse Wedderburn Process
- Applications

Rank Reduction Formula

- An obvious result: (Wedderburn '34)

$$\boxed{A \in R^{m\times n}, x \in R^n, y \in R^m, y^T A x \neq 0}$$
$$+$$
$$\boxed{B := A - \frac{Axy^T A}{y^T A x}}$$
$$\Downarrow$$
$$\boxed{\mathrm{rank}(B) = \mathrm{rank}(A) - 1}$$

- The converse is true! (Egerváry, '60; Householder '75)

$$\boxed{u \in R^m, v \in R^n}$$
$$+$$
$$\boxed{B := A - \sigma^{-1} u v^T}$$
$$+$$
$$\boxed{\mathrm{rank}(B) = \mathrm{rank}(A) - 1}$$
$$\Updownarrow$$
$$\boxed{\exists x \in R^n, y \in R^m \ni u = Ax, v = A^T y, \sigma = y^T A x}$$

- Block form: (Cline & Funderlic, '79)

$$\boxed{U \in R^{m\times k}, R \in R^{k\times k}, V \in R^{n\times k}}$$

$$+$$

$$\boxed{B = A - UR^{-1}V^T}$$

$$+$$

$$\boxed{\text{rank}(B) = \text{rank}(A) - \text{rank}(UR^{-1}V^T)}$$

$$\Updownarrow$$

$$\boxed{\begin{array}{c} \exists X \in R^{n\times k}, Y \in R^{m\times k} \\ \ni U = AX, V = A^TY, R = Y^TAX \end{array}}$$

Repeated Application

- $A_k \neq 0 \Longrightarrow \exists x_k \in R^n, y_k \in R^m \ni \omega_k := y_k^T A_k x_k \neq 0.$
- A sequence of matrices with decreasing ranks:

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

- $\mathrm{rank}(A) = \gamma \Longrightarrow A_{\gamma+1} = 0.$
- Wedderburn decomposition:

$$\begin{aligned} A &= \sum_{k=1}^{\gamma} \omega_k^{-1} A_k x_k y_k^T A_k \\ &:= \Phi \Omega^{-1} \Psi^T \end{aligned}$$

  ⋄ $\Omega := \mathrm{diag}\{\omega_1, \ldots, \omega_\gamma\}.$

  ⋄ $\Phi := [\phi_1, \ldots, \phi_\gamma], \Psi := [\psi_1, \ldots, \psi_\gamma]$

$$\begin{aligned} \phi_k &:= A_k x_k \in R^m \\ \psi_k &:= A_k^T y_k \in R^n. \end{aligned}$$

- Different $\{x_1, \ldots, x_\gamma\}$ and $\{y_1, \ldots, y_\gamma\} \Longrightarrow$ Different decomposition.

An Oblique Projection

- A bilinear form on $R^n \times R^m$:

$$< x, y >:= y^T A x.$$

  ◇ This is NOT an inner product.

- Rewrite the Wedderburn formula:

$$\begin{aligned} Bz &= A\left(z - \frac{< z, y >}{< x, y >}x\right), \\ w^T B &= \left(w^T - \frac{< x, w >}{< x, y >}y^T\right)A. \end{aligned}$$

- Two projectors:

$$\begin{aligned} \mathcal{P}_{A,x,y} &:= I - \frac{xy^T A}{y^T A x} \\ \mathcal{P}'_{A,x,y} &:= I - \frac{Axy^T}{y^T A x}. \end{aligned}$$

- Action of B in terms of A:

$$B = A\mathcal{P}_{A,x,y} = \mathcal{P}'_{A,x,y}A.$$

First Step in Wedderburn Process

- Define:
  - ◇
$$\begin{aligned} u_1 &:= x_1 \\ v_1 &:= y_1 \end{aligned}$$
  - ◇
$$\begin{aligned} u_2 &:= \mathcal{P}x_2 = x_2 - \frac{< x_2, v_1 >}{< u_1, v_1 >} u_1 \\ v_2 &:= (y_2^T \mathcal{P}')^T = y_2 - \frac{< u_1, y_2 >}{< u_1, v_1 >} v_1 \end{aligned}$$

- Results:
$$\begin{aligned} Au_2 &= A_2 x_2 \in \mathcal{R}(A_2) \\ v_2^T A &= y_2^T A_2 \in \mathcal{R}(A_2^T) \\ < u_2, v_1 > &= < u_1, v_2 >= 0 \\ \omega_2 &= y_2^T A_2 x_2 =< u_2, v_2 > . \end{aligned}$$

Wedderburn Process

$$\boxed{\text{rank}(A) = \gamma}$$

$$+$$

$$\boxed{\{x_1, \ldots x_\gamma\} \subset R^n, \{y_1, \ldots y_\gamma\} \subset R^m}$$

$$+$$

$$\boxed{y_k^T A_k x_k \neq 0}$$

$$\Downarrow$$

$$\boxed{u_k := x_k - \Sigma_{i=1}^{k-1} \frac{<x_k, v_i>}{<u_i, v_i>} u_i}$$

$$+$$

$$\boxed{v_k := y_k - \Sigma_{i=1}^{k-1} \frac{<u_i, y_k>}{<u_i, v_i>} v_i}$$

- The process is well defined.
- Bypassing the intermediate A_k:

$$\begin{aligned} Au_k &= A_k x_k, \\ v_k^T A &= y_k^T A_k, \\ \omega_k &= y_k^T A_k x_k =< u_k, v_k > \end{aligned}$$

- Biconjugacy:

$$< u_k, v_j >=< u_j, v_k >= 0.$$

Remarks

- Wedderburn process $\longleftrightarrow$ Gram-Schmidt process:

$$\begin{aligned} x_i \in R^n, y_j \in R^m &\longleftrightarrow \text{Vectors in a single space.} \\ \text{Bilinear form } y^T A x &\longleftrightarrow \text{Standard inner product.} \\ u_i, v_j \text{ biconjugacy} &\longleftrightarrow \text{Orthogonality.} \end{aligned}$$

- Wedderburn decomposition:

$$\begin{aligned} V_k^T A U_k &= \Omega_k, \\ A &= A U_\gamma \Omega_\gamma^{-1} V_\gamma^T A. \end{aligned}$$

  ◇ $\Omega_k := \text{diag}\{\omega_1, \ldots, \omega_k\}$.

- The (1,2)-inverse A^I:

$$A^I = U_\gamma \Omega_\gamma^{-1} V_\gamma^T.$$

  ◇ A nonsingular $\Longrightarrow A^{-1} = U_n \Omega_n^{-1} V_n^T$.

  ◇ An SVD analogue, but U_n and V_n not necessarily orthogonal.

Matrix From of Wedderburn Process

- Rewrite

$$
\begin{aligned}
X_k &= U_k R_k^{(x)}, \\
Y_k &= V_k R_k^{(y)}.
\end{aligned}
$$

◇

$$
R_k^{(x)} = \begin{bmatrix}
1 & \frac{<x_1,v_1>}{<u_1,v_1>} & \cdots & & \frac{<x_j,v_1>}{<u_1,v_1>} & & & \\
0 & 1 & \cdots & & & & & \\
 & \ddots & \ddots & \vdots & \vdots & & & \\
 & & & 1 & \frac{<x_j,v_{j-1}>}{<u_{j-1},v_{j-1}>} & & & \\
 & & & 0 & 1 & & & \\
 & & & & \ddots & \ddots & \vdots & \\
0 & & & & & & 1 & \\
0 & & & & & & 0 & 1
\end{bmatrix}
$$

- Note that $R_k^{(x)}$ depends implicitly on V_k.

Inverse Wedderburn Process

- Think of the Wedderburn process as a function acting on $R^{n\times k} \times R^{m\times k}$ by

$$f(X, Y) := (U, V).$$

  ◇ What is $f^{-1}(U, V)$?

-

$$\boxed{U \in R^{n\times k}, V \in R^{m\times k}}$$

$$+$$

$$\boxed{V^T A U = \Omega \in R^{k\times k} \text{ diagonal}}$$

$$+$$

$$\boxed{X := U R_x, Y := V R_y}$$

$$+$$

$$\boxed{\begin{array}{c} R_x, R_y \\ \text{arbitrary upper triangular with unit diagonal} \end{array}}$$

$$\Downarrow \text{f}$$

$$\boxed{\text{Same } U \text{ and } V}$$

Recover the QR Decomposition

- Suppose $A = QR$.
- Identify $Q^T A R^{-1} = I \Longrightarrow U = R^{-1}, V = Q$.
- Define

$$\begin{aligned} X &:= R^{-1} R_x \\ Y &:= Q R_y. \end{aligned}$$

  ◇ R_x, R_y arbitrary upper triangular with unit diagonal entries.

- Explicit dependence on Q and R (NOT good!)
- Define

$$\begin{aligned} X &:= I \\ Y &:= A. \end{aligned}$$

  ◇ The Wedderburn process produces a matrix U which is upper triangular and a matrix V whose columns are mutually orthogonal. In fact, $V^T V = \Omega$. Furthermore, $AU = V$.

Recover the LR Decomposition

- Suppose $A = LR$.
- Identify $L^{-1}AR^{-1} = I \Longrightarrow U = R^{-1}, V = L^{-T}$.
- The choice

$$\begin{aligned} X &:= R^{-1}R_x \\ Y &:= L^{-T}R_y \end{aligned}$$

for the Wedderburn process will produce the (implicit) LR decomposition of A.

- Define

$$\begin{aligned} X &:= I \\ Y &:= I. \end{aligned}$$

  ◇ The Wedderburn process produces $A = V_n^{-T}\Omega U_n^{-1}$ which is the unique LDM$^{\mathrm{T}}$ decomposition of A.

Relation to SVD

- Freedom in selecting the vectors x_i and y_j.
- Division by $\omega_k = y_k^T A x_k \stackrel{?}{\Longrightarrow}$ Instability.
- Desirable at each stage to choose x_k and y_k so as to

$$\begin{array}{rl} \text{Maximize} & \omega_k = y_k^T A_k x_k = < u_k, v_k > \\ \text{Subject to} & x_k^T x_k = 1,\ y_k^T y_k = 1. \end{array}$$

- The Wedderburn process by meeting this requirement is precisely the SVD of A.

- Lagrange multipliers $\Longrightarrow$
  - ◇ Necessary condition:

$$\begin{aligned} A_k^T \tilde{y}_k &= \tilde{\sigma}_k \tilde{x}_k \\ A_k \tilde{x}_k &= \tilde{\sigma}_k \tilde{y}_k. \end{aligned}$$

  - ◇ Maximal value $= \tilde{\sigma}_k = \|A_k\|_2 =$ Largest singular value of $A_k =$ The kth singular value σ_k of A.
  - ◇ $\tilde{y}_k, \tilde{x}_k =$ the kth left and right singular vectors.
- The Wedderburn process leaves the singular vectors invariant.
- Can X and Y be chosen before the SVD is known? (N0?!)

Other Applications

- Suppose A is symmetric and take $X = Y$. Then $U = V$. The Wedderburn process gives the canonical form of A with respect to congruence.
- With specially selected of X and Y, the Wedderburn process includes the Lanczos algorithm as a special case.
  - ◇ We can explicitly describe X and Y. (Funderlic, November 9, 1993.)
  - ◇ Wedderburn (34) $\Longrightarrow$ Lanczos (50) $\Longrightarrow$ Hestenes and Stiefel (52).
  - ◇ Welcome to the Lanczos Conference (December 12-17, 1993).

Occurrence in BFGS (and DFP)

- Minimize $f : R^n \longrightarrow R$.
  - $\diamond$ Solve $g(x) := \nabla f(x) = 0$.
- Jacobian $G(x)$ of $g(x)$ (Hessian of $f(x)$) is often unavailable or very expensive to compute.
  - $\diamond$ Develop cheap and reasonable approximations to either $G(x)$ or its inverse.
  - $\diamond$ Prefer to preserve symmetry and positive definiteness.

- BFGS — One successful Hessian update:

$$H_+ := H_c + \frac{y_c y_c^T}{y_c^T s_c} - \frac{H_c s_c s_c^T H_c}{s_c^T H_c s_c}.$$

  - ◇ Secant equation:

$$\begin{aligned} H_+ s_c &= y_c \\ s_c &:= x_+ - x_c \\ y_c &:= g(x_+) - g(x_c) \end{aligned}$$

  - ◇ Wedderburn's foumula shows up:

$$H_+ - \frac{H_+ s_c s_c^T H_+}{s_c^T H_+ s_c} = H_c - \frac{H_c s_c s_c^T H_c}{s_c^T H_c s_c}.$$

- What is BFGS doing?
  - ◇ A necessary condition:

$$H_+ \mathcal{P}_{H_+} = H_c \mathcal{P}_{H_c}.$$

  - ◇ Tear and rebuild:

$$H_+ = H_c \underbrace{- \frac{H_c s_c s_c^T H_c}{s_c^T H_c s_c}}_{\text{pruning}} \underbrace{+ \frac{H_+ s_c s_c^T H_+}{s_c^T H_+ s_c}}_{\text{grafting}}.$$