Chapter 8

Inverse Singular Value Problems

- IEP versus ISVP
- Existence question
- A continuous approach
- An iterative method for the IEP
- An iterative method for the ISVP

IEP versus ISVP

- Inverse Eigenvalue Problem (IEP):
  - ◇ Given
    - ▷ Symmetric matrices $A_0, A_1, \ldots, A_n \in R^{n\times n}$;
    - ▷ Real numbers $\lambda_1^* \geq \ldots \geq \lambda_n^*$,
  - ◇ Find
    - ▷ Values of $c := (c_1, \ldots, c_n)^T \in R^n$
    - ▷ Eigenvalues of the matrix

$$A(c) := A_0 + c_1 A_1 + \ldots + c_n A_n$$

are precisely $\lambda_1^*, \ldots, \lambda_n^*$.

- Inverse Singular Value Problem ISVP:
  - ◇ Given
    - ▷ General matrices $B_0, B_1, \ldots, B_n \in R^{m\times n}$, $m \geq n$;
    - ▷ Nonnegative real numbers $\sigma_1^* \geq \ldots \geq \sigma_n^*$,
  - ◇ Find
    - ▷ Values of $c := (c_1, \ldots, c_n)^T \in R^n$
    - ▷ Singular values of the matrix

$$B(c) := B_0 + c_1 B_1 + \ldots + c_n B_n$$

are precisely $\sigma_1^*, \ldots, \sigma_n^*$.

Existence Question

- Not always does the IEP have a solution.
- Inverse Toeplitz Eigenvalue Problem (ITEP)
  - ◇ A special case of the (IEP) where $A_0 = 0$ and $A_k := (A_{ij}^{(k)})$ with
  $$A_{ij}^{(k)} := \begin{cases} 1, & \text{if } |i-j| = k-1; \\ 0, & \text{otherwise.} \end{cases}$$
  - ◇ Symmetric Toeplitz matrices can have *arbitrary* real spectra. (Landau '94, nonconstructive proof by topological degree argument).
- Not aware of any result concerning the existence question for ISVP.

Notation

- $\mathcal{O}(n) :=$ All orthogonal matrices in $R^{n\times n}$;
- $\Sigma = (\Sigma_{ij}) :=$ A "diagonal" matrix in $R^{m\times n}$

$$\Sigma_{ij} := \begin{cases} \sigma_i^*, & \text{if } 1 \le i = j \le n; \\ 0, & \text{otherwise.} \end{cases}$$

- $\mathcal{M}_s(\Sigma) := \{U\Sigma V^T | U \in \mathcal{O}(m), V \in \mathcal{O}(n)\}$
  - ◇ Contains all matrices in $R^{m\times n}$ whose singular values are precisely $\sigma_1^*, \ldots, \sigma_n^*$.
- $\mathcal{B} := \{B(c) | c \in R^n\}$.
- Solving the ISVP ≡ Finding an intersection of the two sets $\mathcal{M}_s(\Sigma)$ and $\mathcal{B}$.

A Continuous Approach

- Assume
  - $\langle B_i, B_j \rangle = \delta_{ij}$ for $1 \le i \le j \le n$.
  - $\langle B_0, B_k \rangle = 0$ for $1 \le k \le n$.
- The projection of X onto the linear subspace spanned by $B_1, \ldots, B_n$:

$$P(X) = \sum_{k=1}^{n} \langle X, B_k \rangle B_k.$$

- The distance from X to the affine subspace $\mathcal{B}$:

$$dist(X, \mathcal{B}) = \|X - (B_0 + P(X))\|.$$

- Define, for any $U \in R^{m \times m}$ and $V \in R^{n \times n}$, a residual matrix:

$$R(U, V) := U\Sigma V^T - (B_0 + P(U\Sigma V^T)).$$

- Consider the optimization problem:

$$\begin{array}{ll} \text{Minimize} & F(U,V) := \frac{1}{2}\|R(U,V)\|^2 \\ \text{Subject to} & (U,V) \in \mathcal{O}(m) \times \mathcal{O}(n). \end{array}$$

Compute the Projected Gradient

- Frobenius inner product on $R^{m\times m} \times R^{n\times n}$:

$$\langle (A_1, B_1), (A_2, B_2)\rangle := \langle A_1, A_2\rangle + \langle B_1, B_2\rangle.$$

- The gradient ∇F may be interpreted as the pair of matrices:

$$\nabla F(U, V) = (R(U, V)V\Sigma^T, R(U, V)^T U\Sigma).$$

- Tangent space can be split:

$$\mathcal{T}_{(U,V)}(\mathcal{O}(m) \times \mathcal{O}(n)) = \mathcal{T}_U\mathcal{O}(m) \times \mathcal{T}_V\mathcal{O}(n).$$

- Projection is easy because:

$$\begin{aligned} R^{n\times n} &= T_V\mathcal{O}(n) \oplus T_V\mathcal{O}(n)^\perp \\ &= V\mathcal{S}(n)^\perp \oplus V\mathcal{S}(n) \end{aligned}$$

- Project the gradient $\nabla F(U,V)$ onto the tangent space $\mathcal{T}_{(U,V)}(\mathcal{O}(m) \times \mathcal{O}(n))$:

$$g(U,V) = \left(\frac{R(U,V)V\Sigma^T U^T - U\Sigma V^T R(U,V)^T}{2}U, \frac{R(U,V)^T U\Sigma V^T - V\Sigma^T U^T R(U,V)}{2}V\right).$$

- Descent vector field:

$$\frac{d(U,V)}{dt} = -g(U,V)$$

defines a steepest descent flow on the manifold $\mathcal{O}(m)\times \mathcal{O}(n)$ for the objective function $F(U,V)$.

The Differential Equation on $\mathcal{M}_s(\Sigma)$

- Define
$$X(t) := U(t)\Sigma V(t)^T.$$
- $X(t)$ satisfies the differential system:
$$\begin{aligned}\frac{dX}{dt} = \; & X\frac{X^T(B_0+P(X))-(B_0+P(X))^T X}{2} \\ & -\frac{X(B_0+P(X))^T-(B_0+P(X))^T X}{2}X.\end{aligned}$$
- $X(t)$ moves on the surface $\mathcal{M}_s(\Sigma)$ in the steepest descent direction to minimize $dist(X(t), \mathcal{B})$.
- This is a continuous method for the ISVP.

Remarks

- No assumption on the multiplicity of singular values is needed.
- Any tangent vector $T(X)$ to $\mathcal{M}_s(\Sigma)$ at a point $X \in \mathcal{M}_s(\Sigma)$ about which a local chart can be defined must be of the form

$$T(X) = XK - HX$$

for some skew symmetric matrices $H \in R^{m\times m}$ and $K \in R^{n\times n}$.

An Iterative Method for IEP

- Assume all eigenvalues $\lambda_1^*, \ldots, \lambda_n^*$ are distinct.
- Consider:
  - ◇ The affine subspace

$$\mathcal{A} := \{A(c) | c \in R^n\}.$$

  - ◇ The isospectral surface

$$\mathcal{M}_e(\Lambda) := \{Q\Lambda Q^T | Q \in \mathcal{O}(n)\}$$

where

$$\Lambda := diag\{\lambda_1^*, \ldots, \lambda_n^*\}.$$

- Any tangent vector $T(X)$ to $\mathcal{M}_e(\Lambda)$ at a point $X \in \mathcal{M}_e(\Lambda)$ must be of the form

$$T(X) = XK - KX$$

for some skew-symmetric matrix $K \in R^{n \times n}$.

A Classical Newton Method

- A function $f : R \longrightarrow R$.
- The scheme:

$$x^{(\nu+1)} = x^{(\nu)} - (f'(x^{(\nu)}))^{-1} f(x^{(\nu)})$$

- The intercept:
  - ◇ The new iterate $x^{(\nu+1)}$ = The x-intercept of the tangent line of the graph of f from $(x^{(\nu)}, f(x^{(\nu)}))$.
- The lifting:
  - ◇ $(x^{(\nu+1)}, f(x^{(\nu+1)}))$ = The natural "lift" of the intercept along the y-axis to the graph of f from which the next tangent line will begin.

An Analogy of the Newton Method

- Think of:
  - ◇ The surface $\mathcal{M}_e(\Lambda)$ as playing the role of the graph of f.
  - ◇ The affine subspace $\mathcal{A}$ as playing the role of the x-axis.
- Given $X^{(\nu)} \in \mathcal{M}_e(\Lambda)$,
  - ◇ There exist a $Q^{(\nu)} \in \mathcal{O}(n)$ such that
$$Q^{(\nu)^T} X^{(\nu)} Q^{(\nu)} = \Lambda.$$
  - ◇ The matrix $X^{(\nu)} + X^{(\nu)}K - KX^{(\nu)}$ with any skew-symmetric matrix K represents a tangent vector to $\mathcal{M}_e(\Lambda)$ emanating from $X^{(\nu)}$.
- Seek an $\mathcal{A}$-intercept $A(c^{(\nu+1)})$ of such a vector with the affine subspace $\mathcal{A}$.
- Lift up the point $A(c^{(\nu+1)}) \in \mathcal{A}$ to a point $X^{(\nu+1)} \in \mathcal{M}_e(\Lambda)$.

Find the Intercept

- Find a skew-symmetric matrix $K^{(\nu)}$ and a vector $c^{(\nu+1)}$ such that

$$X^{(\nu)} + X^{(\nu)}K^{(\nu)} - K^{(\nu)}X^{(\nu)} = A(c^{(\nu+1)}).$$

- Equivalently, find $\tilde{K}^{(\nu)}$ such that

$$\Lambda + \Lambda\tilde{K}^{(\nu)} - \tilde{K}^{(\nu)}\Lambda = {Q^{(\nu)}}^T A(c^{(\nu+1)})Q^{(\nu)}.$$

  ◇ $\tilde{K}^{(\nu)} := {Q^{(\nu)}}^T K^{(\nu)} Q^{(\nu)}$ is skew-symmetric.

- Can find $c^{(\nu)}$ and $K^{(\nu)}$ separately.

- Diagonal elements in the system $\Rightarrow$

$$J^{(\nu)}c^{(\nu+1)} = \lambda^* - b^{(\nu)}.$$

  ◇ Known quantities:

$$\begin{aligned} J_{ij}^{(\nu)} &:= {q_i^{(\nu)}}^T A_j q_i^{(\nu)}, \text{ for } i, j = 1, \ldots, n \\ \lambda^* &:= (\lambda_1^*, \ldots, \lambda_n^*)^T \\ b_i^{(\nu)} &:= {q_i^{(\nu)}}^T A_0 q_i^{(\nu)}, \text{ for } i = 1, \ldots, n \\ q_i^{(\nu)} &= \text{the } i\text{-th column of the matrix } Q^{(\nu)}. \end{aligned}$$

- The vector $c^{(\nu+1)}$ can be solved.

- Off-diagonal elements in the system together with $c^{(\nu+1)} \Rightarrow \tilde{K}^{(\nu)}$ (and, hence, $K^{(\nu)}$):

$$\tilde{K}_{ij}^{(\nu)} = \frac{{q_i^{(\nu)}}^T A(c^{(\nu+1)}) q_j^{(\nu)}}{\lambda_i^* - \lambda_j^*}, \text{ for } 1 \le i < j \le n.$$

Find the Lift-up

- No obvious coordinate axis to follow.
- Solving the IEP $\equiv$ Finding $\mathcal{M}_e(\Lambda) \cap \mathcal{A}$.
- Suppose all the iterations are taking place near a point of intersection. Then

$$X^{(\nu+1)} \approx A(c^{(\nu+1)}).$$

- Also should have

$$A(c^{(\nu+1)}) \approx e^{-K^{(\nu)}} X^{(\nu)} e^{K^{(\nu)}}.$$

- Replace $e^{K^{(\nu)}}$ by the Cayley transform:

$$R := (I + \frac{K^{(\nu)}}{2})(I - \frac{K^{(\nu)}}{2})^{-1} \approx e^{K^{(\nu)}}.$$

- Define

$$X^{(\nu+1)} := R^T X^{(\nu)} R \in \mathcal{M}_e(\Lambda).$$

- The next iteration is ready to begin.

Remarks

- Note that

  $$X^{(\nu+1)} \approx R^T e^{K^{(\nu)}} A(c^{(\nu+1)}) e^{-K^{(\nu)}} R \approx A(c^{(\nu+1)})$$

  represents a lifting of the matrix $A(c^{(\nu+1)})$ from the affine subspace $\mathcal{A}$ to the surface $\mathcal{M}_e(\Lambda)$.

- The above offers a geometrical interpretation of Method III developed by Friedland, Nocedal and Overton (SINUM, 1987).

- Quadratic convergence even for multiple eigenvalues case.

An Iterative Approach for ISVP

- Assume
  - ◇ Matrices $B_0, B_1, \ldots, B_n$ are arbitrary.
  - ◇ All singular values $\sigma_1^*, \ldots, \sigma_n^*$ are positive and distinct.
- Given $X^{(\nu)} \in \mathcal{M}_s(\Sigma)$
  - ◇ There exist $U^{(\nu)} \in \mathcal{O}(m)$ and $V^{(\nu)} \in \mathcal{O}(n)$ such that
    $$U^{(\nu)^T} X^{(\nu)} V^{(\nu)} = \Sigma.$$
  - ◇ Seek a $\mathcal{B}$-intercept $B(c^{(\nu+1)})$ of a line that is tangent to the manifold $\mathcal{M}_s(\Sigma)$ at $X^{(\nu)}$.
  - ◇ Lift the matrix $B(c^{(\nu+1)}) \in \mathcal{B}$ to a point $X^{(\nu+1)} \in \mathcal{M}_s(\Sigma)$.

Find the Intercept

- Find skew-symmetric matrices $H^{(\nu)} \in R^{m\times m}$ and $K^{(\nu)} \in R^{n\times n}$, and a vector $c^{(\nu+1)} \in R^n$ such that

$$X^{(\nu)} + X^{(\nu)}K^{(\nu)} - H^{(\nu)}X^{(\nu)} = B(c^{(\nu+1)})$$

- Equivalently,

$$\Sigma + \Sigma\tilde{K}^{(\nu)} - \tilde{H}^{(\nu)}\Sigma = {U^{(\nu)}}^T B(c^{(\nu+1)})V^{(\nu)}$$

  ◇ Underdetermined skew-symmetric matrices:

$$\begin{aligned}\tilde{H}^{(\nu)} &:= {U^{(\nu)}}^T H^{(\nu)} U^{(\nu)},\\ \tilde{K}^{(\nu)} &:= {V^{(\nu)}}^T K^{(\nu)} V^{(\nu)}.\end{aligned}$$

- Can determine $c^{(\nu+1)}$, $H^{(\nu)}$ and $K^{(\nu)}$ separately.

- Totally $\frac{m(m-1)}{2}+\frac{n(n-1)}{2}+n$ unknowns — the vector $c^{(\nu+1)}$ and the skew matrices $\tilde{H}^{(\nu)}$ and $\tilde{K}^{(\nu)}$.
- Only mn equations.
- Observe: $\tilde{H}_{ij}^{(\nu)}$, $n+1 \le i \ne j \le m$,
  - ◇ $\frac{(m-n)(m-n-1)}{2}$ unknowns.
  - ◇ Locate at the lower right corner of $\tilde{H}^{(\nu)}$.
  - ◇ Are not bound to any equations at all.
  - ◇ Set

$$\tilde{H}_{ij}^{(\nu)} = 0 \text{ for } n+1 \le i \ne j \le m.$$

- Denote

$$W^{(\nu)} := {U^{(\nu)}}^T B(c^{(\nu+1)}) V^{(\nu)}.$$

Then

$$W_{ij}^{(\nu)} = \Sigma_{ij} + \Sigma_{ii}\tilde{K}_{ij}^{(\nu)} - \tilde{H}_{ij}^{(\nu)}\Sigma_{jj},$$

Determine $c^{(\nu+1)}$

- For $1 \le i = j \le n$,

$$J^{(\nu)} c^{(\nu+1)} = \sigma^* - b^{(\nu)}$$

  ◇ Know quantities:

$$\begin{aligned} J_{st}^{(\nu)} &:= {u_s^{(\nu)}}^T B_t v_s^{(\nu)}, \text{ for } 1 \le s, t \le n, \\ \sigma^* &:= (\sigma_1^*, \ldots, \sigma_n^*)^T, \\ b_s^{(\nu)} &:= {u_s^{(\nu)}}^T B_0 v_s^{(\nu)}, \text{ for } 1 \le s \le n. \\ u_s^{(\nu)} &= \text{column vectors of } U^{(\nu)}, \\ v_s^{(\nu)} &= \text{column vectors of } V^{(\nu)}. \end{aligned}$$

- The vector $c^{(\nu+1)}$ is obtained.
- $c^{(\nu+1)} \Rightarrow W^{(\nu)}$.

Determine $H^{(\nu)}$ and $K^{(\nu)}$

- For $n+1 \le i \le m$ and $1 \le j \le n$,

$$\tilde{H}_{ij}^{(\nu)} = -\tilde{H}_{ji}^{(\nu)} = -\frac{W_{ij}^{(\nu)}}{\sigma_j^*}.$$

- For $1 \le i < j \le n$,

$$\begin{aligned} W_{ij}^{(\nu)} &= \Sigma_{ii}\tilde{K}_{ij}^{(\nu)} - \tilde{H}_{ij}^{(\nu)}\Sigma_{jj}, \\ W_{ji}^{(\nu)} &= \Sigma_{jj}\tilde{K}_{ji}^{(\nu)} - \tilde{H}_{ji}^{(\nu)}\Sigma_{ii} \\ &= -\Sigma_{jj}\tilde{K}_{ij}^{(\nu)} + \tilde{H}_{ij}^{(\nu)}\Sigma_{ii}. \end{aligned}$$

Solving for $\tilde{H}_{ij}^{(\nu)}$ and $\tilde{K}_{ij}^{(\nu)} \Rightarrow$

$$\begin{aligned} \tilde{H}_{ij}^{(\nu)} = -\tilde{H}_{ji}^{(\nu)} &= \frac{\sigma_i^* W_{ji}^{(\nu)} + \sigma_j^* W_{ij}^{(\nu)}}{(\sigma_i^*)^2 - (\sigma_j^*)^2}, \\ \tilde{K}_{ij}^{(\nu)} = -\tilde{K}_{ji}^{(\nu)} &= \frac{\sigma_i^* W_{ij}^{(\nu)} + \sigma_j^* W_{ji}^{(\nu)}}{(\sigma_i^*)^2 - (\sigma_j^*)^2}. \end{aligned}$$

- The intercept is now completely found.

Find the Lift-Up

- Define orthogonal matrices

$$\begin{aligned} R &:= (I + \frac{H^{(\nu)}}{2})(I - \frac{H^{(\nu)}}{2})^{-1}, \\ S &:= (I + \frac{K^{(\nu)}}{2})((I - \frac{K^{(\nu)}}{2})^{-1}. \end{aligned}$$

- Define the lifted matrix on $\mathcal{M}_s(\Sigma)$:

$$X^{(\nu+1)} := R^T X^{(\nu)} S.$$

- Observe

$$X^{(\nu+1)} \approx R^T(e^{H^{(\nu)}} B(c^{(\nu+1)}) e^{-K^{(\nu)}})S$$

and

$$\begin{aligned} R^T e^{H^{(\nu)}} &\approx I_m \\ e^{-K^{(\nu)}} S &\approx I_n, \end{aligned}$$

if $\|H^{(\nu)}\|$ and $\|K^{(\nu)}\|$ are small.

- For computation,
  - ◇ Only need orthogonal matrices

$$\begin{aligned} U^{(\nu+1)} &:= R^T U^{(\nu)} \\ V^{(\nu+1)} &:= S^T V^{(\nu)}. \end{aligned}$$

  - ◇ Does not need to form $X^{(\nu+1)}$ explicitly.

Quadratic Convergence

- Measure the discrepancy between $(U^{(\nu)}, V^{(\nu)}) \in R^{m\times m} \times R^{n\times n}$ in the induced Frobenius norm.
- Observe:
  - ◇ Suppose:
    - ▷ The ISVP has an exact solution at c^*.
    - ▷ SVD of $B(c^*) = \hat{U}\Sigma\hat{V}^T$.
  - ◇ Define error matrix::

$$E := (E_1, E_2) := (U - \hat{U}, V - \hat{V}).$$

  - ◇ If $U\hat{U}^T = e^H$ and $V\hat{V}^T = e^K$, then

$$\begin{aligned} U\hat{U}^T &= (E_1 + \hat{U})\hat{U}^T \\ &= I_m + E_1\hat{U}^T \\ &= e^H = I_m + H + O(||H||^2). \end{aligned}$$

  and a similar expression for $V\hat{V}^T$.
  - ◇ Thus,

$$||(H, K)|| = O(||E||).$$

- At the ν-th stage, define

$$E^{(\nu)} := (E_1^{(\nu)}, E_2^{(\nu)}) = (U^{(\nu)} - \hat{U}, V^{(\nu)} - \hat{V}).$$

- How far is ${U^{(\nu)}}^T B(c^*) V^{(\nu)}$ away from Σ?
  - ◇ Write

$$\begin{aligned} &{U^{(\nu)}}^T B(c^*) V^{(\nu)} := e^{-\hat{H}^{(\nu)}} \Sigma e^{\hat{K}^{(\nu)}} \\ &:= ({U^{(\nu)}}^T e^{-H_*^{(\nu)}} U^{(\nu)}) \Sigma ({V^{(\nu)}}^T e^{K_*^{(\nu)}} V^{(\nu)}) \end{aligned}$$

  with

$$\begin{aligned} H_*^{(\nu)} &:= U^{(\nu)} \hat{H}^{(\nu)} {U^{(\nu)}}^T, \\ K_*^{(\nu)} &:= V^{(\nu)} \hat{K}^{(\nu)} {V^{(\nu)}}^T, \end{aligned}$$

  - ◇ Then

$$e^{H_*^{(\nu)}} = U^{(\nu)} \hat{U}^T, e^{K_*^{(\nu)}} = V^{(\nu)} \hat{V}^T.$$

  - ◇ So

$$||(H_*^{(\nu)}, K_*^{(\nu)})|| = O(||E^{(\nu)}||).$$

  - ◇ Norm invariance under orthogonal transformations $\Rightarrow$

$$||(\hat{H}^{(\nu)}, \hat{K}^{(\nu)})|| = O(||E^{(\nu)}||).$$

- Rewrite

$$U^{(\nu)^T} B(c^*) V^{(\nu)} = \Sigma + \Sigma \hat{K}^{(\nu)} - \hat{H}^{(\nu)} \Sigma + O(\|E^{(\nu)}\|^2).$$

- Compare:

$$\begin{aligned} & U^{(\nu)^T} (B(c^*) - B(c^{(\nu+1)})) V^{(\nu)} \\ = \ & \Sigma(\hat{K}^{(\nu)} - \tilde{K}^{(\nu)}) - (\hat{H}^{(\nu)} - \tilde{H}^{(\nu)})\Sigma \\ + \ & O(||E^{(\nu)}||^2). \end{aligned}$$

- Diagonal elements $\Rightarrow$

$$J^{(\nu)}(c^* - c^{(\nu+1)}) = O(||E^{(\nu)}||^2).$$

  ◇ Thus

$$||c^* - c^{(\nu+1)}|| = O(||E^{(\nu)}||^2).$$

- Off-diagonal elements $\Rightarrow$

$$\begin{aligned} ||\hat{H}^{(\nu)} - \tilde{H}^{(\nu)}|| &= O(||E^{(\nu)}||^2), \\ ||\hat{K}^{(\nu)} - \tilde{K}^{(\nu)}|| &= O(||E^{(\nu)}||^2). \end{aligned}$$

  ◇ Therefore,

$$||(\tilde{H}^{(\nu)}, \tilde{K}^{(\nu)})|| = O(||E^{(\nu)}||).$$

- Together,

$$\begin{aligned} ||H^{(\nu)} - H_*^{(\nu)}|| &= O(||E^{(\nu)}||^2), \\ ||K^{(\nu)} - K_*^{(\nu)}|| &= O(||E^{(\nu)}||^2). \end{aligned}$$

- Observe:

$$\begin{aligned} E_1^{(\nu+1)} &:= U^{(\nu+1)} - \hat{U} = R^T U^{(\nu)} - e^{-H_*^{(\nu)}} U^{(\nu)} \\ &= \Big[(I - \frac{H^{(\nu)}}{2}) - (I - H_*^{(\nu)} + O(||H_*^{(\nu)}||^2) \\ &\quad (I + \frac{H^{(\nu)}}{2}) \Big] (I + \frac{H^{(\nu)}}{2})^{-1} U^{(\nu)} \\ &= \Big[H_*^{(\nu)} - H^{(\nu)} + O(||H_*^{(\nu)} H^{(\nu)}|| \\ &\quad + ||H^{(\nu)}||^2) \Big] (I + \frac{H^{(\nu)}}{2})^{-1} U^{(\nu)}. \end{aligned}$$

◇ It is clear now that

$$||E_1^{(\nu+1)}|| = O(||E^{(\nu)}||^2).$$

- A similar argument works for $E_2^{(\nu+1)}$.
- We have proved that

$$||E^{(\nu+1)}|| = O(||E^{(\nu)}||^2).$$

Multiple Singular Values

- Previous definition in finding the $\mathcal{B}$-intercept of a tangent line of $\mathcal{M}_s(\Sigma)$ allows
  - ◇ No zero singular values.
  - ◇ No multiple singular values.
- Now assume
  - ◇ All singular values are positive.
  - ◇ Only the first singular value σ_1^* is multiple, with multiplicity p.

- Observe:
  - ◇ All formulas work, except
    - ▷ For $1 \leq i < j \leq p$, only know

$$W_{ij}^{(\nu)} + W_{ji}^{(\nu)} = 0.$$

    - ▷ No values for $\tilde{H}_{ij}^{(\nu)}$ and $\tilde{K}_{ij}^{(\nu)}$ can be determined.
    - ▷ Additional $q := \frac{p(p-1)}{2}$ equations for the vector $c^{(\nu+1)}$ arise.
- Multiple singular values gives rise to an overdetermined system for $c^{(\nu+1)}$.
  - ◇ Tangent lines from $\mathcal{M}_s(\Sigma)$ may not intercept the affine subspace $\mathcal{B}$ at all.
  - ◇ The ISVP needs to be modified.

Modified ISVP

- Given
  - ◇ Positive values $\sigma_1^* = \ldots = \sigma_p^* > \sigma_{p+1}^* > \ldots > \sigma_{n-q}^*$,
- Find
  - ◇ Real values of $c_1, \ldots, c_n$,
  - ◇ The $n - q$ largest singular values of the matrix matrix $B(c)$ are $\sigma_1^*, \ldots, \sigma_{n-q}^*$.

Find the Intercept

- Use the equation

$$\hat{\Sigma} + \hat{\Sigma}\tilde{K}^{(\nu)} - \tilde{H}^{(\nu)}\hat{\Sigma} = {U^{(\nu)}}^T B(c^{(\nu+1)}) V^{(\nu)}$$

  to find the $\mathcal{B}$-intercept where

  ◇ The diagonal matrix

$$\hat{\Sigma} := diag\{\sigma_1^*, \ldots, \sigma_{n-q}^*, \hat{\sigma}_{n-q+1}, \ldots, \hat{\sigma}_n\}$$

  ◇ Additional singular values $\hat{\sigma}_{n-q+1}, \ldots, \hat{\sigma}_n$ are free parameters.

The Algorithm

Given $U^{(\nu)} \in \mathcal{O}(m)$ and $V^{(\nu)} \in \mathcal{O}(n)$,

- Solve for $c^{(\nu+1)}$ from the system of equations:

$$\sum_{k=1}^{n} \left({u_i^{(\nu)}}^T B_k v_i^{(\nu)} \right) c_k^{(\nu+1} = \sigma_i^* - {u_i^{(\nu)}}^T B_0 v_i^{(\nu)},$$

for $i = 1, \ldots, n - q$

$$\sum_{k=1}^{n} \left({u_s^{(\nu)}}^T B_k v_t^{(\nu)} + {u_t^{(\nu)}}^T B_k v_s^{(\nu)} \right) c_k^{(\nu+1)} =$$

$$-{u_s^{(\nu)}}^T B_0 v_t^{(\nu)} - {u_t^{(\nu)}}^T B_0 v_s^{(\nu)},$$

for $1 \leq s < t \leq p$.

- Define $\hat{\sigma}_k^{(\nu)}$ by

$$\hat{\sigma}_k^{(\nu)} := \begin{cases} \sigma_k^*, & \text{if } 1 \leq k \leq n - q; \\ {u_k^{(\nu)}}^T B(c^{(\nu+1)}) v_k^{(\nu)}, & \text{if } n - q < k \leq n \end{cases}$$

- Once $c^{(\nu+1)}$ is determined, calculate $W^{(\nu)}$.

- Define skew symmetric matrices $\tilde{K}^{(\nu)}$ and $\tilde{H}^{(\nu)}$:
  - ◇ For $1 \leq i < j \leq p$, the equation to be satisfied is
$$W_{ij}^{(\nu)} = \hat{\sigma}_i^{(\nu)} \tilde{K}_{ij}^{(\nu)} - \tilde{H}_{ij}^{(\nu)} \hat{\sigma}_j^{(\nu)}.$$
    - ▷ Many ways to define $\tilde{K}_{ij}^{(\nu)}$ and $\tilde{H}_{ij}^{(\nu)}$.
    - ▷ Set $\tilde{K}_{ij}^{(\nu)} \equiv 0$ for $1 \leq i < j \leq p$.
  - ◇ $\tilde{K}^{(\nu)}$ is defined by
$$\tilde{K}_{ij}^{(\nu)} := \begin{cases} \frac{\hat{\sigma}_i^{(\nu)} W_{ij}^{(\nu)} + \hat{\sigma}_j^{(\nu)} W_{ji}^{(\nu)}}{(\hat{\sigma}_i^{(\nu)})^2 - (\hat{\sigma}_j^{(\nu)})^2}, & \text{if } 1 \leq i < j \leq n;\ p < j; \\ 0, & \text{if } 1 \leq i < j \leq p \end{cases}$$
  - ◇ $\tilde{H}^{(\nu)}$ is defined by
$$\tilde{H}_{ij}^{(\nu)} := \begin{cases} -\frac{W_{ij}^{(\nu)}}{\hat{\sigma}_j^{(\nu)}}, & \text{if } 1 \leq i < j \leq p; \\ -\frac{W_{ij}^{(\nu)}}{\hat{\sigma}_j^{(\nu)}}, & \text{if } n+1 \leq i \leq m;\ 1 \leq j \leq n; \\ \frac{\hat{\sigma}_i^{(\nu)} W_{ji}^{(\nu)} + \hat{\sigma}_j^{(\nu)} W_{ij}^{(\nu)}}{(\hat{\sigma}_i^{(\nu)})^2 - (\hat{\sigma}_j^{(\nu)})^2}, & \text{if } 1 \leq i < j \leq n;\ p < j; \\ 0, & \text{if } n+1 \leq i \neq j \leq m. \end{cases}$$
- Once $\tilde{H}^{(\nu)}$ and $\tilde{K}^{(\nu)}$ are determined, proceed the lifting in the same way as for the ISVP.

Remarks

- No longer on a fixed manifold $\mathcal{M}_s(\Sigma)$ since $\hat{\Sigma}$ is changed per step.
- The algorithm for multiple singular value case converges quadratically.

Zero Singular Value

- Zero singular value $\Rightarrow$ rank deficiency.
- Finding a lower rank matrix in a generic affine subspace $\mathcal{B}$ is intuitively a more difficult problem.
- More likely the ISVP does not have a solution.
- Consider the simplest case where $\sigma_1^* > \ldots > \sigma_{n-1}^* > \sigma_n^* = 0$.
  - ◇ Except for $\tilde{H}_{in}$ (and $\tilde{H}_{ni}$), $i = n+1, \ldots, m$, all other quantities including $c^{(\nu+1)}$ are well-defined.
  - ◇ It is necessary that
$$W_{in}^{(\nu)} = 0 \text{ for } i = n+1, \ldots, m.$$
  - ◇ If the necessary condition fails, then no tangent line of $\mathcal{M}_s(\Sigma)$ from the current iterate $X^{(\nu)}$ will intersect the affine subspace $\mathcal{B}$.

Example of the Continuous Approach

- Integrator — Subroutine ODE (Shampine et al, '75).
  - ◇ ABSERR and RELERR $= 10^{-12}$.
  - ◇ Output values examined at interval of 10.
- Two consecutive output points differ by less than $10^{-10} \Rightarrow$ Convergence.
- Stable equilibrium point is not necessarily a solution to the ISVP.
- Change to different initial value $X(0)$ if necessary.

Example of the Iterative Approach

- Easy implementation by MATLAB.
  - ◇ Consider the case when $m = 5$ and $n = 4$.
  - ◇ Randomly generated basis matrices by the Gaussian distribution.
- Numerical experiment meant solely to examine the quadratic convergence.
  - ◇ Randomly generate a vector $c^{\#} \in R^4$.
  - ◇ Singular values of $B(c^{\#})$ used as the prescribed singular values.
  - ◇ Perturb each entry of $c^{\#}$ by a uniform distribution between -1 and 1.
  - ◇ Use the perturbed vector as the initial guess.

Observations

- The limit point c^* is not necessary the same as the original vector $c^{\#}$.
- Singular values of $B(c^*)$ do agree with those of $B(c^{\#})$.
- Differences between singular values of $B(c^{(\nu)})$ and $B(c^*)$ are measured in the 2-norm.
- Quadratic convergence is observed.

Example of Multiple Singular Values

- Construction of an example is not trivial.
  - ◇ Same basis matrices as before.
  - ◇ Assume $p = 2$.
  - ◇ Prescribed singular values $\sigma^* = (5, 5, 2)^T$.
  - ◇ Initial guess of $c^{(0)}$ is searched by trials
- The order of singular values could be altered.
  - ◇ The value 5 is no longer the largest singular value.
  - ◇ Unless the initial guess $c^{(0)}$ is close enough to an exact solution c^*, no reason to expect that the algorithm will preserve the ordering.
  - ◇ Once convergence occurs, then σ^* must be part of the singular values of the final matrix.
- At the initial stage the convergence is slow, but eventually the rate is picked up and becomes quadratic.