Chapter 5

Simultaneous Reduction Problems

- Overview
- Reduction by orthogonal similarity
- Gradient flow for orthogonal similarity
- Nearest commuting pair problem
- Reduction by orthogonal equivalence
- Descent flow for orthogonal equivalence
- Nearest normal matrix problem

Overview

What is the simplest form to which a family of matrices depending smoothly on the parameters can be reduced by a change of coordinates depending smoothly on the parameters?

- Arnold '88

- Two types of transformations:
  - ◇ Similarity transformation $Q^T AQ$.
  - ◇ Equivalence transformation $Q^T AZ$.

- Applications:
  - ◇ Computation of eigenvalues, generalized eigenvalues and singular values.
  - ◇ Simultaneous reduction of more than one matrices:
    - ▷ System identification in control.
    - ▷ Multibody oscillations in mechanics.
  - ◇ Matrix canonical forms.
- Significance:
  - ◇ Very few theories or numerical methods are available.
  - ◇ Projected gradient method offers an easy but versatile reduction procedure.

Reduction by Orthogonal Similarity

- The problem:
  - ◇ Matrices A_i and subspaces V_i in $R^{n\times n}$, $i = 1, \ldots, p$, are given.
  - ◇ Find the best reduction of A_i to an element in V_i through orthogonal similarity transformation.
- The idea:
  - ◇ $P_i(X)$ = projection of X onto V_i.
  - ◇ Simultaneously reduce the norm of all residuals:

$$\alpha_i(Q) := Q^T A_i Q - P_i(Q^T A_i Q).$$

- Formulation:

$$\begin{array}{rcl} \text{Minimize} & F(Q) & := \frac{1}{2} \sum_{i=1}^{k} \|\alpha_i(Q)\|^2 \\ \text{Subject to} & Q & \in \mathcal{O}(n). \end{array}$$

Gradient Flow under Similarity

- The gradient:

$$\nabla F(X) = \sum_{i=1}^{k} (A_i^T X \alpha_i(X) + A_i X \alpha_i^T(X)).$$

- The projected gradient on $\mathcal{O}(n)$:

$$\begin{aligned} g(Q) &= \frac{Q}{2}\{Q^T \nabla F(X) - \nabla F(X)^T Q\} \\ &= \frac{Q}{2} \sum_{i=1}^{k} \left([Q^T A_i^T Q, \alpha_i(Q)] + [Q^T A_i Q, \alpha_i^T(Q)]\right). \end{aligned}$$

- The descent flow:

$$\begin{aligned} \frac{dX_i}{dt} &= \left[X_i, \sum_{j=1}^{p} \frac{[X_j, P_j^T(X_j)] - [X_j, P_j^T(X_j)]^T}{2}\right] \\ X_i(0) &= A_i \end{aligned}$$

$\diamond$ $X_i := Q^T A Q$.

Examples

- A continuous Jacobi algorithm for symmetric eigenvalue problem:

$$\begin{aligned} \frac{dX}{dt} &= [X, [X, diag(X)]] \\ X(0) &= A_1 \text{ (symmetric)} \end{aligned}$$

  - ◇ V_1 = subspace of all diagonal matrices.
  - ◇ Diagonal matrices are the only stable equilibrium points for the system.
  - ◇ Any of these isospectral diagonal matrices is a global minimizer.

- A similar idea for general eigenvalue problem?

$$\begin{aligned} \frac{dX}{dt} &= \left[X, \frac{[X, P_1^T(X)] - [X, P_1^T(X)]^T}{2}\right] \\ X(0) &= A_1 \text{ (general).} \end{aligned}$$

  ◇ V_1 = subspace of upper triangular matrices.

$$\begin{bmatrix} 1.0000 & 3.0000 & 5.0000 & 7.0000 \\ -3.0000 & 1.0000 & 2.0000 & 4.0000 \\ 0.0000 & 0.0000 & 3.0000 & 5.0000 \\ 0.0000 & 0.0000 & 0.0000 & 4.0000 \end{bmatrix}$$

$$\downarrow$$

$$\begin{bmatrix} 2.2500 & 3.3497 & 3.1713 & 2.8209 \\ -0.3506 & 2.2500 & 8.0562 & 6.1551 \\ 0.6247 & -0.8432 & 2.2500 & 3.2105 \\ -0.0846 & 0.2727 & -0.3360 & 2.2500 \end{bmatrix}$$

    ▷ Upper quasi-triangular matrices are not necessarily stable.

  ◇ V_1 = subspace of structure B matrices, i.e., block upper triangular matrices with all diagonal blocks 2×2 except possibly the last one, which is 1×1.

    ▷ ω-limit set contains only a singleton.

The Nearest Commuting-Pair Problem

- The problem:
  - ◇ Given a pair of symmetric matrices (A_1, A_2).
  - ◇ Determine how far (A_1, A_2) is away from a symmetric pair (E_1, E_2) satisfying $E_1E_2 = E_2E_1$.
- Symmetric and commuting pairs can be simultaneously diagonalized by orthogonal similarity transformation.
- Reformulation:

$$\sum_{i=1}^{2} \|E_i - A_i\|^2 = \sum_{i=1}^{2} \|D_i - Q^T A_i Q\|^2$$

  - ◇ $D_i = Q^T E_i Q$ is diagonal.
- The descent flow:

$$\begin{aligned} \frac{dX_i}{dt} &= \left[X_i, \sum_{j=1}^{2}[X_j, diag(X_j)]\right] \\ X_i(0) &= A_i, \quad i = 1, 2 \end{aligned}$$

Reduction by Orthogonal Equivalence

- The problem:
  - ◇ Matrices A_i and subspaces V_i in $R^{m\times n}$, $i = 1, \ldots, p$, are given.
  - ◇ Find the best reduction of A_i to an element in V_i through orthogonal equivalence transformation .
- Formulation:

$$\begin{array}{rl} \text{Minimize} & F(Q,Z) := \frac{1}{2}\sum_{i=1}^{k} \|\beta_i(Q,Z)\|^2 \\ \text{Subject to} & Q \in O(m) \\ & Z \in O(n). \end{array}$$

  - ◇ Residuals:

$$\beta_i(Q,Z) := Q^T A_i Z - P(Q^t A_i Z).$$

Gradient Flow under Equivalence

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

$$\langle (X_1, Y_1), (X_2, Y_2)\rangle_P := \langle X_1, X_2\rangle + \langle Y_1, Y_2\rangle.$$

- The gradient:

$$\nabla F(X,Y) = \left(\sum_{i=1}^{k} A_i Y \alpha_i^T(X,Y), \sum_{i=1}^{k} A_i^T X \alpha_i(X,Y)\right).$$

- The projected gradient on $\mathcal{O}(m) \times \mathcal{O}(n)$:

$$g(Q,Z) = \left(\frac{Q}{2}\sum_{i=1}^{k}(Q^T A_i Z \alpha_i^T(Q,Z) - \alpha_i(Q,Z) Z^T A_i^T Q), \frac{Z}{2}\sum_{i=1}^{k}(Z^T A_i^T Q \alpha_i(Q,Z) - \alpha_i^T(Q,Z) Q^T A_i Z)\right).$$

- The decent flow:

$$\begin{aligned}\frac{dX_i}{dt} &= \sum_{j=1}^{p}\left\{X_i\frac{X_j^T P_j(X_j) - P_j^T(X_j)X_j}{2}\right.\\ &\qquad \left. + \frac{P_j(X_j)X_j^T - X_j P_j^T(X_j)}{2}X_i\right\}\\ X_i(0) &= A_i\end{aligned}$$

◇ $X_i := Q^T A_i Z$.

Example

- A continuous Jacobi algorithm for the singular value problem:

$$\frac{dX}{dt} = \frac{1}{2}\left\{X\left(X^T(diagX) - (diagX)^T X\right) + \left((diagX)X^T - X(diagX)^T\right)X.\right\}$$

  - ◇ V_1 = subspace of diagonal matrices in $R^{m\times n}$.
  - ◇ Diagonal matrices (of the singular values) are the only stable equilibrium points.

Nearest Normal Matrix Problem

- The problem:
  - ◇ Determine the closest normal matrix to a given square complex matrix A.
  - ◇ Normal matrices can be diagonalized by unitary similarity transformation. Thus

$$\begin{array}{rl} \text{Minimize} & F(U,D) := \frac{1}{2}\|A - UDU^*\|^2 \\ \text{Subject to} & U \in \mathcal{U}(n) \\ & D \in \mathcal{D}(n). \end{array}$$

- Reformulation:
  - ◇ Consider U and D as independent variables:

$$\|A - Z\|^2 = \|U^*AU - D\|^2.$$

  - ◇ At global minimum,

$$\begin{array}{rl} \text{Minimize} & F(U) = \frac{1}{2}\|U^*AU - diag(U^*AU)\|^2 \\ \text{Subject to} & U^*U = I. \end{array}$$

- Identify $C^{n\times n} \equiv R^{n\times n} \times R^{n\times n}$ to take advantage of earlier techniques:
  - ⋄ $Z = (\Re Z, \Im Z)$.
  - ⋄ Inner product:
  $$\langle X, Y\rangle_C := \langle \Re X, \Re Y\rangle + \langle \Im X, \Im Y\rangle.$$
  - ⋄ The project gradient of F onto the manifold $\mathcal{U}(n)$ is given by:
  $$\begin{aligned} g(U) \; = \; & \frac{U}{2}\left\{[diag(U^*AU), U^*A^*U] \right. \\ & \left. - [diag(U^*AU), U^*A^*U]^*\right\}. \end{aligned}$$
- New characterization: Let $W := U^*AU$. Then necessary (sufficient) conditions for $U \in \mathcal{U}(n)$ to be a local minimizer are that:
  - ⋄ The matrix $[diag(W), W^*]$ is Hermitian.
  - ⋄ The quadratic form
  $$\langle [diag(W), K] - diag([W, K]), [W, K]\rangle_C$$
  is nonnegative (positive) for every skew Hermitian matrix K.

- The descent flow:

$$\frac{dU}{dt} = Uk(W)$$
$$\frac{dW}{dt} = [W, k(W)]$$

  - ◇ $k(W) := \frac{1}{2}[W, diag(W^*)] - [W, diag(W^*)]^*$.
  - ◇ $W(t) := U(t)^* A U(t)$.
  - ◇ $Z := \tilde{U} diag(\tilde{W}) \tilde{U}^*$ (where $\tilde{}$ denotes a limit point of the flow) is a putative nearest normal matrix.

- Advantages:
  - ◇ The necessary conditions are derived without reference to Lagrange multipliers.
  - ◇ Computation is easier. No shift, phase, or rotation angles are needed as in the Jacobi algorithm. (Goldstine et al. '59, Ruhe '87).