Chapter 7

Spectrally Constrained Approximation

- Overview
- Reformulation
- Least Squares Approximation
- Toeplitz Inverse Eigenvalue Problem
- Eigenvalue Computation
- Simultaneous Reduction
- Nearest Normal Matrix Approximation

Overview

- Least squares approximations for various types of real and symmetric matrices subject to spectral constraints share a common structure.
- The projected gradient can be formulated explicitly.
- A descent flow can be followed numerically.
- The procedure can be extended to approximating general matrices subject to singular value constraints.
- Notation:

$$\begin{aligned}
\mathcal{S}(n) &:= \{\text{All real symmetric matrices}\} \\
\mathcal{O}(n) &:= \{\text{All real orthogonal matrices}\} \\
||X|| &:= \text{ Frobenius matrix norm of } X \\
\Lambda &:= \text{ A given matrix in } \mathcal{S}(n) \\
M(\Lambda) &:= \{Q^T \Lambda Q | Q \in \mathcal{O}(n)\} \\
\mathcal{V} &:= \text{ A single matrix or a subspace in } \mathcal{S}(n) \\
P(X) &:= \text{ The projection of } X \text{ into } \mathcal{V}
\end{aligned}$$

$$\begin{aligned}
\Sigma &:= \text{ A given general matrix in } R^{m \times n} \\
W(\Sigma) &:= \{Q_1 \Sigma Q_2 | Q_1 \in \mathcal{O}(m), Q_2 \in \mathcal{O}(n)\} \\
\mathcal{U} &:= \text{ A single matrix or a subspace in } \mathbb{R}^{m \times n}
\end{aligned}$$

Spectrally Constrained Problem

$$\begin{aligned} \text{Minimize } F(X) &:= \frac{1}{2}||X - P(X)||^2 \\ \text{Subject to } X &\in M(\Lambda) \end{aligned}$$

- Special cases:
  - ◇ Problem A: Given a symmetric matrix, find its least squares approximation with prescribed spectrum.
  - ◇ Problem B: Construct a symmetric Toeplitz matrix that has a prescribed set of eigenvalues.
  - ◇ Problem C: Find the spectrum of a given a symmetric matrix.

Singular-Value Constrained Problem

$$\begin{aligned} \text{Minimize } F(X) &:= \frac{1}{2}||X - P(X)||^2 \\ \text{Subject to } X &\in W(\Sigma) \end{aligned}$$

- Special cases:
  - ◇ Problem D: Given a general real $m \times n$ matrix, find its least square approximation that has a prescribed set of singular values.
  - ◇ Problem E: Construct a general real $m \times n$ matrix, find its singular values.

Reformulation

- Idea:
  1. $X \in M(\Lambda)$ satisfies the spectral constraint.
  2. $P(X) \in V$ has the desirable structure in V.
  3. Minimize the undesirable part $\|X - P(X)\|$.
- Working with the parameter Q is easier:

$$
\begin{aligned}
\text{Minimize } F(Q) \; &:= \; \frac{1}{2}\Big\langle Q^T \Lambda Q - P(Q^T \Lambda Q), \\
& \qquad Q^T \Lambda Q - P(Q^T \Lambda Q)\Big\rangle \\
\text{Subject to } Q^T Q \; &= \; I
\end{aligned}
$$

  ◇ $\langle A, B\rangle = \operatorname{trace}(AB^T)$ is the Frobenius inner product.

Feasible Set $O(n)$ & Gradient of F

- The set $O(n)$ is a regular surface.
- The tangent space of $O(n)$ at any orthogonal matrix Q is given by

$$T_Q O(n) = QK(n)$$

  where

$$K(n) = \{\text{All skew-symmetric matrices}\}.$$

- The normal space of $O(n)$ at any orthogonal matrix Q is given by

$$N_Q O(n) = QS(n).$$

- The Fréchet Derivative of F at a general matrix A acting on B:

$$F'(A)B = 2\langle \Lambda A(A^T \Lambda A - P(A^T \Lambda A)), B\rangle.$$

- The gradient of F at a general matrix A:

$$\nabla F(A) = 2\Lambda A(A^T \Lambda A - P(A^T \Lambda A)).$$

The Projected Gradient

- A splitting of $R^{n\times n}$:

$$\begin{aligned} R^{n\times n} &= T_Q O(n) + N_Q O(n) \\ &= QK(n) + QS(n). \end{aligned}$$

- A unique orthogonal splitting of $X \in R^{n\times n}$:

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

- The projection of $\nabla F(Q)$ into the tangent space:

$$\begin{aligned} g(Q) &= Q\left\{\frac{1}{2}(Q^T \nabla F(Q) - \nabla F(Q)^T Q)\right\} \\ &= Q[P(Q^T \Lambda Q), Q^T \Lambda Q]. \end{aligned}$$

An Isospectral Descent Flow

- A descent flow on the manifold $O(n)$:

$$\frac{dQ}{dt} = Q[Q^T \Lambda Q, P(Q^T \Lambda Q)].$$

- A descent flow on the manifold $M(\Lambda)$:

$$\begin{aligned} \frac{dX}{dt} &= \frac{dQ^T}{dt}\Lambda Q + Q^T \Lambda \frac{dQ}{dt} \\ &= [X, \underbrace{[X, P(X)]}_{k(X)}]. \end{aligned}$$

- The entire concept can be obtained by utilizing the Riemannian geometry on the Lie group $O(n)$.

The Second Order Derivative

- Extend the projected gradient g to the function

$$G(Z) := Z[P(Z^T \Lambda Z), Z^T \Lambda Z]$$

for general matrix Z.

- The Fréchet derivative of G:

$$\begin{aligned} G'(Z)H \quad &= H[P(Z^T \Lambda Z), Z^T \Lambda Z] \\ &+ Z[P(Z^T \Lambda Z), Z^T \Lambda H + H^T \Lambda Z] \\ &+ Z[P'(Z^T \Lambda Z)(Z^T \Lambda H + H^T \Lambda Z), Z^T \Lambda Z]. \end{aligned}$$

- The projected Hessian at a critical point $X = Q^T \Lambda Q$ for the tangent vector QK with $K \in K(n)$:

$$\begin{aligned} &\langle G'(Q)QK, QK \rangle = \\ &\langle [P(X), K] - P'(X)[X, K], [X, K] \rangle. \end{aligned}$$

Least Squares Approximation

- Let the given matrix be $\hat{A}$ and $\Lambda := \operatorname{diag}\{\lambda_1, \ldots, \lambda_n\}$. The projection is $P(X) = \hat{A}$.
- The projected gradient is given by:

$$g(Q) = Q[\hat{A}, Q^T \Lambda Q].$$

- The descent flow is given by the IVP:

$$\begin{aligned} \frac{dX}{dt} &= [[\hat{A}, X], X] \\ X(0) &= \Lambda. \end{aligned}$$

Sorting Property

- Assume the given eigenvalues are $\lambda_1 > \ldots > \lambda_n$.
- Assume the eigenvalues of $\hat{A}$ are $\mu_1 > \ldots > \mu_n$.
- Assume Q is a critical point on $O(n)$ and define

$$\begin{aligned} X &:= Q^T \Lambda Q \\ E &:= Q\hat{A}Q^T. \end{aligned}$$

- The first order condition $[\hat{A}, X] = 0$ implies E must be a diagonal matrix. Hence, the diagonals of E must be a permutation of $\mu_1, \ldots, \mu_n$.
- The second order derivative is reduced to

$$\begin{aligned} \langle G'(Q)QK, QK \rangle &= \langle [\hat{A}, K], [X, K] \rangle \\ &= \langle E\hat{K} - \hat{K}E, \Lambda\hat{K} - \hat{K}\Lambda \rangle \\ &= 2\sum_{i<j} (\lambda_i - \lambda_j)(e_i - e_j)\hat{k}_{ij}^2. \end{aligned}$$

Wielandt-Hoffman Theorem

- We have shown that if a matrix Q is optimal, then the columns $q_1, \ldots, q_n$ of Q^T must be the normalized eigenvectors of $\hat{A}$ corresponding respectively to $\mu_1, \ldots, \mu_n$. The solution to Problem A is unique and is given by

$$X = \lambda_1 q_1 {q_1}^T + \ldots + \lambda_n q_n {q_n}^T.$$

- Let A and $A + E$ be symmetric matrices with eigenvalues $\mu_1 > \ldots \mu_n$ and $\lambda_1 > \ldots > \lambda_n$, respectively. Then

$$\sum_{i=1}^{n} (\lambda_i - \mu_i)^2 \leq ||E||^2.$$

Toeplitz Inverse Eigenvalue Problem

- Let $\mathcal{T}$ be the subspace of all symmetric Toeplitz matrices and $\Lambda := \operatorname{diag}\{\lambda_1, \ldots, \lambda_n\}$.
- The subspace $\mathcal{T}$ has a natural orthogonal basis, say $E_1, \ldots, E_n$. So the projection of any matrix X is given by

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

- The projected gradient is given by:

$$g(Q) = Q[P(Q^T \Lambda Q), Q^T \Lambda Q].$$

- The descent flow is given by the IVP:

$$\begin{aligned} \frac{dX}{dt} &= [[P(X), X], X] \\ X(0) &= \text{any thing on } M(\Lambda) \text{ but diagonal matrices.} \end{aligned}$$

- Open Question: With an arbitrary structured affined subspace $\mathcal{V}$ (See the IEP with Prescribed Entries), characterize the critical points of the descent flow.

Toeplitz Annihilator

- To stay on the surface $\mathcal{M}(\Lambda)$, a differential equation must take the form
$$\frac{dX}{dt} = [X, k(X)]$$
where $k : \mathcal{S}(n) \longrightarrow \mathcal{S}(n)^{\perp}$.

- Require k to be a linear Toeplitz annihilator:
  - ◇ $k(X) = 0$ if and only if $X \in \mathcal{T}$.

- What is the idea?
  - ◇ Suppose all elements in Λ are distinct.
  - ◇ $[X, k(X)] = 0$ if and only if $k(X)$ is a polynomial of X.
  - ◇ $k(X) \in \mathcal{S}(n) \cap \mathcal{S}(n)^{\perp} = \{0\}$.
  - ◇ $||X(t)|| = ||\Lambda||$ for all $t \in R$.
  - ◇ A bounded flow on a compact set must have a non-empty ω-limit set.

- Can such a k be defined?
  - ◇ The simpliest choice:

$$k_{ij} := \begin{cases} x_{i+1,j} - x_{i,j-1}, & \text{if } 1 \leq i < j \leq n \\ 0, & \text{if } 1 \leq i = j \leq n \\ x_{i,j-1} - x_{i+1,j}, & \text{if } 1 \leq j < i \leq n \end{cases}$$

- Open Question: Starting with the unique centro-symmetric Jacobi matrix as the initial value, must the annihilator flow converge? [119]

Eigenvalue Computation

- Let V be the subspace of all diagonal matrices and $\Lambda = X_0$ be the matrix whose eigenvalues are to be found.
- The objective of Problem C is the same as that of the Jacobi method, i.e., to minimize the off-diagonal elements.
- The descent flow is given by the IVP:

$$\begin{aligned} \frac{dX}{dt} &= [[\mathrm{diag}(X), X], X] \\ X(0) &= X_0. \end{aligned}$$

- The necessary condition for X to be critical is

$$[\mathrm{diag}(X), X] = 0.$$

Simultaneous Reduction

- Simultaneous reduction of real matrices by either orthogonal similarity or orthogonal equivalence transformation is hard [64].
  - ◇ Little is known in both theory and practice on how reduction for more than two matrices.
  - ◇ The project gradient method based on the Jacobi idea can be formulated.
- Simultaneous reduction flow:

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

- Nearest normal matrix problem [64]

$$\begin{aligned} \frac{dW}{dt} &= \left[W, \frac{1}{2}[W, diag(W^*)] - [W, diag(W^*)]^* \right] \\ W(0) &= A. \end{aligned}$$