Chapter 4

Spectrally Constrained Problems

- Overview
- General framework
- Special cases
- Reformulation
- Geometry of $\mathcal{O}(n)$
- Projected gradient
- Isospectral descent flow
- Second order condition
- Least squares matrix approximation
- Inverse eigenvalue problem
- Symmetric eigenvalue problem
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 general matrices subject to singular value constraints.
General Framework

Minimize \( F(X) := \frac{1}{2} \|X - P(X)\|^2 \)
Subject to \( X \in M(\Lambda) \)

\[ S(n) := \{\text{All real symmetric matrices}\} \]
\[ O(n) := \{\text{All real orthogonal matrices}\} \]
\[ \|X\| := \text{Frobenius matrix norm of } X \]
\[ \Lambda := \text{A given matrix in } S(n) \]
\[ M(\Lambda) := \{Q^T \Lambda Q | Q \in O(n)\} \]
\[ V := \text{A single matrix or a subspace in } S(n) \]
\[ P(X) := \text{The projection of } X \text{ into } V \]
Special Cases

• Problem A:
  ◊ Find the least squares approximation of a given symmetric matrix subject to a prescribed set of eigenvalues.

• Problem B:
  ◊ Construct a symmetric Toeplitz matrix that has a prescribed set of eigenvalues.

• Problem C:
  ◊ Calculate the eigenvalue of a given symmetric matrix.
Reformulation

- Rewrite the problem in terms of the coordinate variable $Q$:

\[
\text{Minimize } F(Q) := \frac{1}{2} \langle Q^T \Lambda Q - P(Q^T \Lambda Q),
Q^T \Lambda Q - P(Q^T \Lambda Q) \rangle
\]

Subject to $Q^T Q = I$

\[\langle A, B \rangle = \text{trace}(AB^T)\] is the Frobenius inner product.
Geometry of $O(n)$

- The set $O(n)$ is a regular surface.
- The tangent space of $O(n)$ at any orthogonal matrix $Q$ is given by
  \[ T_QO(n) = QK(n) \]
  \[ \diamond \]
  \[ K(n) = \{ \text{All skew-symmetric matrices} \}. \]
- The normal space of $O(n)$ at any orthogonal matrix $Q$ is given by
  \[ N_QO(n) = QS(n). \]
Projected Gradient

- 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 $R^{n \times n}$:
  \[ R^{n \times n} = T_{QO}(n) + N_{QO}(n) = QK(n) + QS(n). \]
- Any $X \in R^{n \times n}$ has a unique orthogonal splitting:
  \[ X = Q\left\{\frac{1}{2}(Q^T X - X^T Q) + \frac{1}{2}(Q^T X + X^T Q)\right\}. \]
- The gradient $\nabla F(Q)$ can be projected into the tangent space easily:
  \[ g(Q) = Q\left\{\frac{1}{2}(Q^T \nabla F(Q) - \nabla F(Q)^T Q)\right\} = Q[P(Q^T \Lambda Q), Q^T \Lambda Q]. \]
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 $M(\Lambda)$:
  \[ \frac{dX}{dt} = \frac{dQ^T}{dt} \Lambda Q + Q^T \Lambda \frac{dQ}{dt} = [X, [X, P(X)]] . \]
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 $G$:

$$G'(Z)H = H[P(Z^T \Lambda Z), Z^T \Lambda Z] + Z[P(Z^T \Lambda Z), Z^T \Lambda H + H^T \Lambda Z] + Z[P'(Z^T \Lambda Z)(Z^T \Lambda H + H^T \Lambda Z), Z^T \Lambda Z].$$

- The projected Hessian at a critical point $X = QT \Lambda Q$ for the tangent vector $QK$ with $K \in K(n)$ is described explicitly by the quadratic form:

$$\langle G'(Q)QK, QK \rangle =$$

$$\langle [P(X), K] - P'(X)[X, K], [X, K] \rangle.$$
Least Squares Matrix Approximation

- Given a symmetric matrix $N$, find its least squares approximation whose eigenvalues are $\{\lambda_1, \ldots, \lambda_n\}$.

- **Setup:**
  - $\Lambda := \text{diag}\{\lambda_1, \ldots, \lambda_n\}$.
  - The projection is $P(X) = N$.

- The projected gradient:
  $$g(Q) = Q[N, Q^T \Lambda Q].$$

- The descent flow:
  $$\frac{dX}{dt} = [X, [X, N]]$$
  $$X(0) = \Lambda.$$
Second Order Condition

- **Assume**
  - ♦ The given eigenvalues are $\lambda_1 > \ldots > \lambda_n$.
  - ♦ The eigenvalues of $N$ are $\mu_1 > \ldots > \mu_n$.
  - ♦ $Q$ is a critical point on $O(n)$ and define
    \[
    X := Q^T \Lambda Q \\
    E := QNQ^T.
    \]

- **The first order condition:**
  \[
  [N, X] = 0
  \]
  - ♦ $E$ must be a diagonal matrix.
  - ♦ $E$ must be a permutation of $\mu_1, \ldots, \mu_n$.

- **The projected Hessian:**
  \[
  \langle G'(Q)QK, QK \rangle = \langle [N, K], [X, K] \rangle \\
  = \langle E\hat{K} - \hat{K}E, \Lambda\hat{K} - \hat{K}\Lambda \rangle \\
  = 2 \sum_{i<j} (\lambda_i - \lambda_j)(e_i - e_j)k_{ij}^2.
  \]
Significance

- If a matrix $Q$ is optimal, then:
  - Columns of $Q^T = [q_1, \ldots, q_n]$ must be the normalized eigenvectors of $N$ corresponding in the order to $\mu_1, \ldots, \mu_n$.
  - The solution is unique.
  - The solution is given by
    \[ X = \lambda_1 q_1 q_1^T + \ldots + \lambda_n q_n q_n^T. \]

- We have reproved the Wielandt-Hoffman theorem.

- The dynamics in the problem enjoys a special sorting property.
  - Can be applied to data matching and a variety of combinatorial optimization problems, including the LP problem.
Inverse Toeplitz Eigenvalue Problem

- Construct a symmetric Toeplitz matrix whose eigenvalues are $\{\lambda_1, \ldots, \lambda_n\}$

- Setup:
  - The set $\mathcal{T}$ of all symmetric Toeplitz matrices forms a linear subspace with a natural basis $E_1, \ldots, E_n$.
  - $\Lambda := \text{diag}\{\lambda_1, \ldots, \lambda_n\}$.
  - The projection of any matrix $X$ is easy:
    $$P(X) = \sum_{i=1}^{n} \langle X, E_i \rangle E_i$$

- The projected gradient:
  $$g(Q) = Q[P(Q^T \Lambda Q), Q^T \Lambda Q]$$

- The descent flow:
  $$\frac{dX}{dt} = [X, [X, P(X)]]$$
  $$X(0) = \text{Anything but diagonals in } M(\Lambda)$$
Significance

- The descent flow approach offers a globally convergent method for solving the inverse Toeplitz eigenvalue problem.

- A stable critical point may not be Toeplitz.

- The second order condition has not be analyzed yet. Further work is needed.
Symmetric Eigenvalue Problem

- Setup:
  - $\Lambda = X_0$, the matrix whose eigenvalues are to be found.
  - $V$ = the subspace of all diagonal matrices.
  - $P(X) = \text{diag}(X)$.

- The objective is the same as that of the Jacobi method, i.e., to minimize the off-diagonal elements.

- The descent flow:
  \[
  \frac{dX}{dt} = [X, [X, \text{diag}(X)]] \\
  X(0) = X_0.
  \]
• Let $X$ be a critical point. Then
  
  ◦ If $X$ is a diagonal matrix, then $X$ is a global minimizer.
  
  ◦ If $X$ is not a diagonal matrix but $\text{diag}(X)$ is a scalar matrix, then $X$ is a global maximizer.
  
  ◦ If $X$ is not a diagonal matrix and $\text{diag}(X)$ is not a scalar matrix, then $X$ is a saddle point.