Extremal-trace problems

Maximizing and minimizing quadratic forms

given $n \times n$ symmetric matrix A

$$\begin{aligned} \text{maximize}: \quad & x^\mathsf{T} A x \\ \text{subject to}: \quad & \|x\| = 1 \end{aligned}$$

eigenvalue decomposition of A:

$$\sum_{i=1}^{n} \lambda_i q_i q_i^\mathsf{T}$$

- $\lambda_1 \geq \cdots \geq \lambda_n$ eigenvalues of A

- $q_1, \ldots, q_n \in \mathbb{R}^n$ orthonormal eigenvectors

- solution: $x = q_1$

- optimal value: λ_1

to minimize:

- solution: $x = q_n$

- optimal value: λ_n

Maximizing and minimizing sums of quadratic forms

$$\begin{array}{ll}\text{maximize}: & \sum_{i=1}^{k} x_i^\mathsf{T} A x_i \\ \text{subject to}: & \|x_i\| = 1 \\ & x_i^\mathsf{T} x_j = 0 \quad i \neq j \end{array}$$

compact representation:

$$\begin{array}{ll}\text{maximize}: & \mathbf{Tr}(X^\mathsf{T} A X) \\ \text{subject to}: & X^\mathsf{T} X = I \end{array}$$

▶ solution: $x_1 = q_1, \ldots, x_k = q_k$

▶ optimal value: $\lambda_1 + \cdots + \lambda_k$

to minimize:

▶ solution: $x_1 = q_n, \ldots, x_k = q_{n-k+1}$

▶ optimal value: $\lambda_n + \cdots + \lambda_{n-k+1}$

Maximizing the input gain

given $A \in \mathbb{R}^{m \times n}$

$$\begin{array}{ll} \text{maximize :} & \sum_{i=1}^{k} \|Ax_i\|^2 \\ \text{subject to :} & \|x_i\| = 1, \\ & x_i^\mathsf{T} x_j = 0 \quad i \neq j \end{array}$$

extremal-trace formulation:

$$\begin{array}{ll} \text{maximize :} & \mathsf{Tr}(X^\mathsf{T} A^T A X) \\ \text{subject to :} & X^\mathsf{T} X = I \end{array}$$

with variable $X \in \mathbb{R}^{n \times k}$

equivalent ways of expressing the solution:

▶ $X = \begin{bmatrix} q_1 & \cdots & q_k \end{bmatrix}$, where $q_1, \ldots, q_k$ are eigenvectors of $A^\mathsf{T} A$

▶ $X = \begin{bmatrix} v_1 & \cdots & v_k \end{bmatrix}$, where $v_1, \ldots, v_k$ are right singular vectors of A

Application: worst-case earthquake

simple model of building during earthquake:

$$x(t+1) = Ax(t) + Bu(t), \qquad x(0) = 0,$$
$$y(t) = Cx(t)$$

- ▶ u is ground displacement

- ▶ y is displacement of top of building

- ▶ problem data:

$$A = \begin{bmatrix} 0.9 & 0.5 \\ -0.5 & 0.7 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \qquad \text{and} \qquad C = \begin{bmatrix} 1 & 2 \end{bmatrix}$$

- ▶ assumptions about input:

$$\sum_{t=0}^{T_u} \|u(t)\|^2 \le 1 \qquad \text{and} \qquad u(t) = 0, \qquad t > T_u$$

where $T_u = 50$

Application: worst-case earthquake

find worst-case earthquake according to criterion

$$\sum_{t=1}^{T}\|y(t)\|^2,$$

where $T = 100$:

$$\text{maximize}: \quad \|Mu\|^2$$
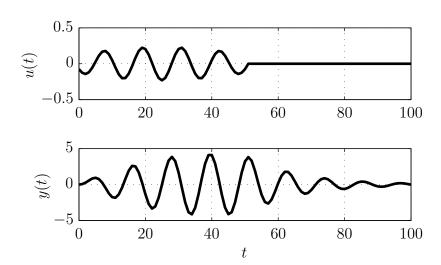$$\text{subject to}: \quad \|u\|^2 \leq 1,$$

where

$$M = \begin{bmatrix} CB & 0 & 0 & \cdots & 0 \\ CAB & CB & 0 & \cdots & 0 \\ CA^2 & CAB & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ CA^{T-1}B & CA^{T-2}B & CA^{T-3}B & \cdots & CA^{T-T_u-1}B \end{bmatrix}$$

Application: worst-case earthquake

solution: first right singular vector of M

Application: worst-case earthquake
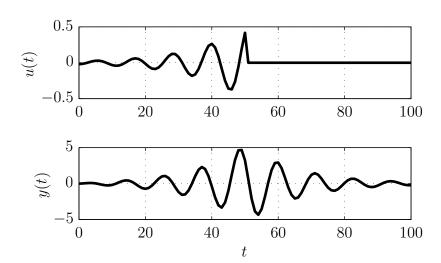
find worst-case earthquake according to criterion $\|y(T)\|^2$:

$$\text{maximize :} \quad \|\tilde{M}u\|^2$$
$$\text{subject to :} \quad \|u\|^2 \leq 1,$$

where

$$\tilde{M} = \begin{bmatrix} CA^{T-1}B & CA^{T-2}B & CA^{T-3}B & \cdots & CA^{T-T_u-1}B \end{bmatrix}$$

Application: worst-case earthquake

solution: first right singular vector of $\tilde{M}$

Minimizing the projection error

given $A \in \mathbb{R}^{m \times n}$

$$\begin{array}{ll} \text{minimize :} & \sum_{j=1}^{n} \|(I - \sum_{i=1}^{k} x_i x_i^{\mathsf{T}}) a_j\|^2 \\ \text{subject to :} & \|x_i\| = 1, \\ & x_i^{\mathsf{T}} x_j = 0 \quad i \neq j \end{array}$$

where a_j is jth column of A

equivalently, maximize the projection power

$$\begin{array}{ll} \text{maximize :} & \sum_{i=1}^{k} \sum_{j=1}^{n} (x_i^{\mathsf{T}} a_j)^2 \\ \text{subject to :} & \|x_i\| = 1, \\ & x_i^{\mathsf{T}} x_j = 0 \quad i \neq j \end{array}$$
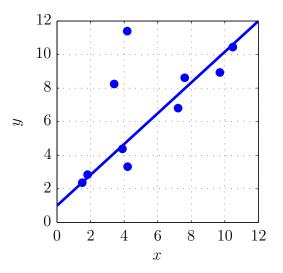
Minimizing the projection error

extremal-trace formulation:

$$\begin{aligned} \text{maximize} : &\quad \mathbf{Tr}(X^\mathsf{T} A A^\mathsf{T} X) \\ \text{subject to} : &\quad X^\mathsf{T} X = I \end{aligned}$$

with variable $X \in \mathbb{R}^{n \times k}$

equivalent ways of expressing the solution:

- $X = \begin{bmatrix} q_1 & \cdots & q_k \end{bmatrix}$, where $q_1, \ldots, q_k$ are eigenvectors of $A A^\mathsf{T}$

- $X = \begin{bmatrix} u_1 & \cdots & u_k \end{bmatrix}$, where $u_1, \ldots, u_k$ are left singular vectors of A

Application: total least squares

data points $(x_i, y_i) \in \mathbb{R}^2$ for $i = 1, \ldots, n$: $y_i = a(x_i + \delta_i) + b + \epsilon_i$, $a = b = 1$
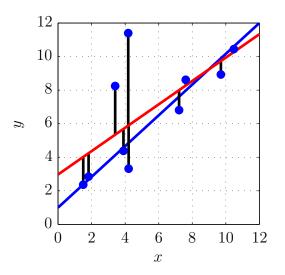
Application: least squares

least squares chooses $\hat{a}$ and $\hat{b}$ to minimize

$$\sum_{i=1}^{n}(\hat{a}x_i + \hat{b} - y_i)^2$$

solution:

$$\begin{bmatrix} \hat{a}_{\mathrm{ls}} \\ \hat{b}_{\mathrm{ls}} \end{bmatrix} = \begin{bmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix}^{\dagger} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

Application: total least squares

least squares: $\hat{a}_{\mathrm{ls}} = 0.70, \hat{b}_{\mathrm{ls}} = 2.97$; relative error in $(\hat{a}_{\mathrm{ls}}, \hat{b}_{\mathrm{ls}}) = 1.99$

Application: total least squares

total least squares chooses x_0, y_0, and d to minimize

$$\sum_{i=1}^{n} \min_{t_i \in \mathbb{R}} \left\| \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} + t_i d - \begin{bmatrix} x_i \\ y_i \end{bmatrix} \right\|^2$$

subject to $\|d\| = 1$

solution:

$$x_0 = \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i, \qquad y_0 = \bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i$$

$$d = \text{first left singular vector of } \begin{bmatrix} x_1 - \bar{x} & \cdots & x_n - \bar{x} \\ y_1 - \bar{y} & \cdots & y_n - \bar{y} \end{bmatrix}$$

recover slope and intercept:

$$\hat{a}_{\text{tls}} = \frac{d_2}{d_1} \qquad \text{and} \qquad \hat{b}_{\text{tls}} = \bar{y} - \hat{a}_{\text{tls}} \bar{x}$$

Application: total least squares

total least squares: $\hat{a}_{\mathrm{tls}} = 1.07, \hat{b}_{\mathrm{tls}} = 0.98$; relative error in $(\hat{a}_{\mathrm{tls}}, \hat{b}_{\mathrm{tls}}) = 0.07$