Inverse Euler-Lagrange Method for Constrained Quadratic Optimization

Some Humans

07.06.2025

Abstract

This note addresses the optimization of a quadratic form $F(\varphi) = \langle A\varphi, A\varphi\rangle$ on $V = \mathbb{R}^n$, subject to the constraint $G(\varphi) = \langle B\varphi, B\varphi\rangle = C \in \mathbb{R}$, where $A, B \in \mathbb{R}^{n\times n}$. Using the Euler-Lagrange method, we derive the generalized eigenvalue problem $A^*A\varphi_0 = \lambda B^*B\varphi_0$, whose solutions minimize (maximize) F on S. We prove that negative eigenvalues correspond to local minima and positive eigenvalues to maxima in case if B is positive definite.

1 Problem Statement

We are given a vector space V and a symmetric inner product $\langle\cdot,\cdot\rangle$ as well as quadratic forms

$$\begin{aligned} &F : \mathbb{R}^n \to \mathbb{R}, \varphi \mapsto F(\varphi) = \langle A\varphi, A\varphi\rangle, A \in \mathbb{R}^{n\times n}, \\ &G : \mathbb{R}^n \to \mathbb{R}, \varphi \mapsto G(\varphi) = \langle B\varphi, B\varphi\rangle, B \in \mathbb{R}^{n\times n} \end{aligned} \tag{1}$$

and a constant $C \in \mathbb{R}$.

The task is to find φ_0 s.t. $\varphi_0 = \min\{F(\varphi) : \varphi \in \mathbb{R}^n \text{ satisfying } G(\varphi) = C\}$. This reduces to solving the following generalized eigenvalue problem:

$$A^*A\varphi_0 = \lambda B^*B\varphi_0, \tag{2}$$

which we discuss in this short note.

1.1 Notations

We write $S := \{\varphi \in V : G(\varphi) = C\}$ for the surface we optimize F on.

2 Derivation

We start by defining the Lagrangian $\mathcal{L}(\varphi, \lambda) := F(\varphi) - \lambda(G(\varphi) - C)$. Finding its minimum is equivalent to solving the problem above because

$$\nabla\mathcal{L}(\varphi, \lambda) = \begin{pmatrix} \nabla_\varphi[F(\varphi) - \lambda G(\varphi)] \\ \nabla_\lambda[F(\varphi) - \lambda(G(\varphi) - C)] \end{pmatrix} = \begin{pmatrix} \nabla_\varphi F(\varphi) - \lambda\nabla_\varphi G(\varphi) \\ -G(\varphi) + C \end{pmatrix}, \tag{3}$$

thus it being zero for some φ_0, ensures both extremum of $F(\varphi_0)$ on S and $G(\varphi_0) = C$.

2.1 Gradients

One observes that

$$\nabla_\varphi[F(\varphi_0) - \lambda G(\varphi_0)] = 0 \iff DF(\varphi_0)[h] = \lambda DG(\varphi_0)[h] \text{ for all } h \in V. \tag{4}$$

By definition,

$$
\begin{aligned}
DF(\varphi_0)[h] &= \lim_{t\to 0}\frac{F(\varphi_0+th)-F(\varphi_0)}{t} = \lim_{t\to 0}(th+2\langle A\varphi_0, Ah\rangle) = 2\langle A\varphi_0, Ah\rangle \\
&= 2\langle A^*A\varphi_0, h\rangle
\end{aligned}
\tag{5}
$$

By analogy,

$$
DG(\varphi_0)[h] = 2\langle B\varphi_0, Bh\rangle = 2\langle B^*B\varphi_0, h\rangle. \tag{6}
$$

2.2 Final Steps

Using (5) and (6) and setting (3) to zero, we obtain

$$
\langle A^*A\varphi_0, h\rangle = \lambda\langle B^*B\varphi_0, h\rangle \text{ for all } h \in V, \tag{7}
$$

which is equivalent to solving the following generalized eigenvalue problem:

$$
A^*A\varphi_0 = \lambda B^*B\varphi_0 \tag{8}
$$

Theorem 2.2.1: A solution to the GEP (8) is a local minimum of F on S if $\lambda < 0$ and maximum if $\lambda > 0$.

Proof:

Let φ_0, λ be a solution to (8). Because of the considerations above, $\varphi_0 \in S$. Then take any $h \in V \setminus \{0\}$, s.t. $\varphi_0 + h \in S$. Note that

$$
\begin{aligned}
C = G(\varphi_0+h) &= \underbrace{G(\varphi_0)}_{=C} + G(h) + 2\langle B\varphi_0, Bh\rangle \\
&= C + G(h) + 2\langle B\varphi_0, Bh\rangle.
\end{aligned}
$$

Thus,

$$
G(h) = -2\langle B\varphi_0, Bh\rangle > 0.
$$

And in the end,

$$
\begin{aligned}
F(\varphi_0+h) &= F(\varphi_0) + F(h) + 2\langle A\varphi_0, Ah\rangle \\
&= F(\varphi_0) + F(h) + 2\lambda\langle B\varphi_0, Bh\rangle \\
&= F(\varphi_0) + \underbrace{F(h)}_{\geq 0}\underbrace{-\lambda G(h)}_{\geq 0 \text{ if } \lambda \leq 0} \geq F(\varphi_0).
\end{aligned}
$$

□