MATH 7502 project: Constrained Least Squares Problems

Team Members: Ziyi Zhou(45792708) Xuerong Wang(45540534) Shibo Liu(44935768) Yuting Sun(45044528)

1. Linearly constrained least squares

the (linearly) constrained least squares problem (**CLS**) is

minimize $\|Ax - b\|^2$

subject to $Cx = d$

$\|Ax - b\|^2$ is the objective function and $\mathrm{Cx} = \mathrm{d}$ are equality constraints, x is an n-vector, A is a m × n matrix, b is an m-vector, C is a p × n matrix, and d is a p-vector.

$\hat{\mathrm{x}}$ is a solution of CLS if $\mathrm{C}\hat{\mathrm{x}} = \mathrm{d}$ and $\|\mathrm{A}\hat{\mathrm{x}} - \mathrm{b}\| \le \|Ax - b\|^2$ holds for any n-vector x that satisfies $\mathrm{Cx} = \mathrm{d}$.

2. Methods of solving the constrained least squares problem

3.1. KKT equations

optimality conditions in matrix-vector form: $2(A^T A)\hat{x} - 2A^T \mathrm{b} + C^T \mathrm{z} = 0, \mathrm{C}\hat{\mathrm{x}} = \mathrm{d}$, put these together to get KKT conditions $\begin{bmatrix} 2(A^T A) & C^T \\ C & 0 \end{bmatrix} \begin{bmatrix} \hat{\mathrm{x}} \\ \hat{\mathrm{z}} \end{bmatrix} = \begin{bmatrix} 2A^T \mathrm{b} \\ \mathrm{d} \end{bmatrix}$, then we have

$$\begin{bmatrix} \hat{\mathrm{x}} \\ \hat{\mathrm{z}} \end{bmatrix} = \begin{bmatrix} 2(A^T A) & C^T \\ C & 0 \end{bmatrix}^{-1} \begin{bmatrix} 2A^T \mathrm{b} \\ \mathrm{d} \end{bmatrix}$$

3.2. QR factorization

We can solve the constrained least squares problem via QR factorization.

3. Least norm problem (An important special case of the constrained least squares problem)

the least norm problem (LN) is

minimize $\|x\|^2$

subject to $c_i^T x = d_i$, i=1, . . ., p

How to get the optimality conditions for Least norm problem

3.1. **Form Lagrangian function, with Lagrange multipliers** (Derivatives of the Cost) $\boldsymbol{z_1, \ldots, z_p}$. $\mathrm{L}(\mathrm{x}, \mathrm{z}) = \mathrm{f}(\mathrm{x}) + z_1(c_1{}^T x - d_1) + \cdots + z_p(c_p{}^T x - d_p)$

3.1.1. Short introduction to Lagrange multipliers

In mathematical optimization, the method of Lagrange multipliers is a strategy for finding the local maxima and minima of a function subject to equality constraints. The basic idea is to convert a constrained problem into a form such that the derivative test of an unconstrained problem can still be applied.

A simple example that helps you understand

Suppose we want to find the minimum values of

$$\mathrm{f}(\mathrm{x}, \mathrm{y}) = \mathrm{x}^2\mathrm{y}$$

with the condition$\mathrm{x}^2 + \mathrm{y}^2 = 1$

As there is just a single constraint, we will use only one multiplier, say λ.

$$\mathrm{g}(\mathrm{x}, \mathrm{y}) = \mathrm{x}^2 + \mathrm{y}^2 - 1$$

$$\Phi(\mathrm{x}, \mathrm{y}, \lambda) = \mathrm{f}(\mathrm{x}, \mathrm{y}) + \lambda \mathrm{g}(\mathrm{x}, \mathrm{y}) = \mathrm{x}^2\mathrm{y} + \lambda(\mathrm{x}^2 + \mathrm{y}^2 - 1)$$

Now we can calculate the gradient of $\Phi(\mathrm{x}, \mathrm{y}, \lambda)$, and set it to be 0, then we have:

$$\begin{cases} 2\mathrm{xy} + 2\lambda\mathrm{x} = 0 \\ \mathrm{x}^2 + 2\lambda\mathrm{y} = 0 \\ \mathrm{x}^2 + \mathrm{y}^2 - 1 = 0 \end{cases}$$

The minimum value is one of the solutions of the following system of equations

3.2. **Optimality conditions** are $\frac{\partial \mathrm{L}}{\partial x_i}(\hat{\mathrm{x}}, \hat{\mathrm{z}}) = 2\sum_{j=1}^{n}(A^T A)_{ij}\, \hat{x}_j - 2(\mathrm{A}^T b)_i + \sum_{j=1}^{p} \hat{z}_j (c_j)_i = 0$, i = 1, . . ., n, $\frac{\partial \mathrm{L}}{\partial z_i}(\hat{\mathrm{x}}, \hat{\mathrm{z}}) = c_i{}^T x - d_i = 0$, i = 1, . . ., p.

4. Applications

4.1. Linear quadratic control

The theory of optimal control is concerned with operating a dynamic system at minimum cost. The case where the system dynamics are described by a set of linear differential equations and the cost is described by a quadratic function is called the LQ problem.

4.2. Linear quadratic state estimation

For the linear quadratic state estimation, what they can do is that calculating the input noise and measure noise to be applied in navigation system and all guidance like global positioning system.