SI152: Numerical Optimization

Lecture 14: Quadratic Programming

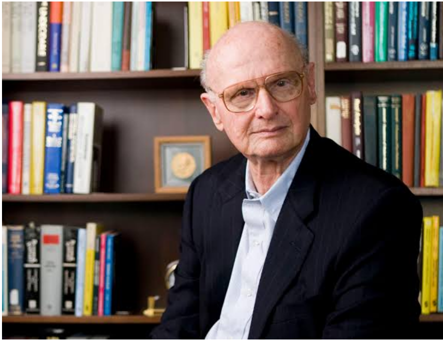
Hao Wang
Email: haw309@gmail.com

ShanghaiTech University

November 25, 2024

1. Example: Markowitz Model
2. Active-set Method
3. Interior Point Method

Harry Markowitz (1927-2023). Markowitz's seminal 1952 paper, "Portfolio Selection," introduced the concept of risk diversification and laid the foundation for modern investment strategies. His introduction of the efficient frontier and mean-variance optimization revolutionized financial economics. Markowitz was awarded the Nobel Prize in Economic Sciences in 1990.



$R_p = \sum_{i=1}^n w_i R_i$, $\sum_{i=1}^n w_i = 1, w \geq 0$.

1. **Expected Return**: for a portfolio with n assets, the expected return $\mathbb{E}[R_p]$ is given by weighted sum of the expected returns of the individual assets:

$$\mathbb{E}[R_p] = \sum_{i=1}^{n} w_i \mu_i,$$

- w_i is the weight of asset i in the portfolio,
- $\mu_i := \mathbb{E}[R_i]$ is the expected return of asset i.

2. **Portfolio Variance**: The risk of the portfolio is measured by the covariance of its returns

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \dots & \sigma_{nn} \end{pmatrix},$$

where each entry $\sigma_{ij} = \mathrm{Cov}(R_i, R_j)$ is the covariance between R_i and R_j

$$\sigma_{ij} = \mathbb{E}[(R_i - \mathbb{E}[R_i])(R_j - \mathbb{E}[R_j])].$$

Given a vector of portfolio weights $\mathbf{w}$, the portfolio risk is given by: $\mathbf{w}^T \Sigma \mathbf{w}$.

The goal of the Markowitz model is to construct a portfolio with the minimum possible risk for a given expected return. The optimization problem can be formulated as:

$$\min_{\mathbf{w}} \quad \mathbf{w}^T \Sigma \mathbf{w} \quad \text{s.t. } \mathbf{1}^T \mathbf{w} = 1, \mathbf{w} \geq 0 \quad \text{and} \quad \mathbf{w}^T \mu \geq r,$$

$$\min_{\mathbf{w}} \quad \mathbf{w}^T \Sigma \mathbf{w} - \lambda \mathbf{w}^T \mu \quad \text{s.t. } \mathbf{1}^T \mathbf{w} = 1, \mathbf{w} \geq 0,$$

$$\min_{\mathbf{w}} -\mathbf{w}^T \mu \qquad \text{s.t. } \mathbf{1}^T \mathbf{w} = 1, \mathbf{w} \geq 0, \mathbf{w}^T \Sigma \mathbf{w} \leq s.$$

where:

- $\mathbf{w}$ is the vector of portfolio weights $(w_1, w_2, \ldots, w_n)$,
- Σ is the covariance matrix of asset returns,
- $\mathbf{1}$ is a vector of ones, ensuring that the weights sum to 1,
- $\mathbf{r}$ is the vector of expected returns for each asset.

Outline

1. Example: Markowitz Model
2. Active-set Method
3. Interior Point Method

Consider the equality constrained quadratic optimization problem

$$\min_x \; f(x) = g^T x + \frac{1}{2} x^T H x$$
$$\text{s.t. } Ax + b = 0$$

The Lagrangian is

$$L(x, \lambda) = g^T x + \frac{1}{2} x^T H x + \lambda^T (Ax - b)$$

so necessary optimality conditions are

$$\begin{bmatrix} g + Hx + A^T \lambda \\ Ax + b \end{bmatrix} = 0 \implies \begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} = - \begin{bmatrix} g \\ b \end{bmatrix}$$

Necessary conditions above are sufficient as long as

$$d^T H d \geq 0 \;\text{ for all } d \text{ such that } Ad = 0$$

Consider the equality constrained quadratic optimization problem

$$\begin{aligned} \min_x \; & f(x) = g^T x + \frac{1}{2} x^T H x \\ \text{s.t.} \quad & A_{\mathcal{E}} x + b_{\mathcal{E}} = 0, \\ & A_{\mathcal{I}} x + b_{\mathcal{I}} \leq 0. \end{aligned}$$

CQ holds, so the KKT conditions are

$$\begin{aligned} g + Hx + A_{\mathcal{E}}^T \lambda_{\mathcal{E}} + A_{\mathcal{I}}^T \lambda_{\mathcal{I}} &= 0 \\ A_{\mathcal{E}} x + b_{\mathcal{E}} &= 0 \\ A_{\mathcal{I}} x + b_{\mathcal{I}} &= 0 \\ \lambda_{\mathcal{I}} &\geq 0 \\ \lambda_{\mathcal{I}} \cdot (A_{\mathcal{I}} x + b_{\mathcal{I}}) &= 0. \end{aligned}$$

If an optimal active-set $\mathcal{A}_*$ (i.e., a set of inequalities satisfied as equalities at a solution) is known in advance, then a solution x_* can be found as a solution to

$$\min_x \; g^T x + \frac{1}{2} x^T H x$$
$$\text{s.t. } A_i x + b_i = 0, \; i \in \mathcal{E} \cup \mathcal{A}(x_*),$$

i.e., a solution to

$$\begin{aligned} g + Hx + \sum_{i \in \mathcal{E} \cup \mathcal{A}(x_*)} A_i^T \lambda_i &= 0 \\ A_i x + b_i &= 0, i \in \mathcal{E} \cup \mathcal{A}(x_*). \end{aligned}$$

Suppose we have an iterate x^k and a guess $\mathcal{A}^k$ of an optimal active set. Compute d^k as the solution to the subproblem

$$\min_d \; g^T(x^k+d) + \frac{1}{2}(x^k+d)^T H(x^k+d)$$
$$\text{s.t. } A_i(x^k+d) + b_i = 0, i \in \mathcal{E} \cup \mathcal{A}^k.$$

- If $x^k + d^k$ is feasible, then set $x_{k+1} \leftarrow x^k + d^k$ and let $\mathcal{A}_{k+1} \leftarrow \mathcal{A}^k$
- Else, set $x_{k+1} \leftarrow x^k + \alpha^k d^k$, where α^k is the largest value such that x_{k+1} satisfies all constraints. Let $\mathcal{A}_{k+1}$ be the set of constraints active at x_{k+1}.

Continue this process until $d^k = 0$ for some k...

Eventually, we obtain a solution (x^k, λ^k) of the KKT conditions

$$\begin{aligned} g + Hx + \sum_{i \in \mathcal{E} \cup \mathcal{A}^k} A_i^T \lambda_i =& 0 \\ A_i x + b_i =& 0, i \in \mathcal{E} \cup \mathcal{A}^k. \end{aligned}$$

(with setting $\lambda_i = 0$ for $i \in \mathcal{I} \setminus \mathcal{A}^k$). The KKT conditions for the quadratic problem

$$\begin{aligned} g + Hx + A_{\mathcal{E}}^T \lambda_{\mathcal{E}} + A_{\mathcal{I}}^T \lambda_{\mathcal{I}} &= 0 \\ A_{\mathcal{E}} x + b_{\mathcal{E}} &= 0 \\ A_{\mathcal{I}} x + b_{\mathcal{I}} &= 0 \\ \lambda_{\mathcal{I}} &\geq 0 \\ \lambda_{\mathcal{I}} \cdot (A_{\mathcal{I}} x + b_{\mathcal{I}}) &= 0. \end{aligned}$$

will be satisfied as long as x^k is feasible and $\lambda_i^k \geq 0, i \in \mathcal{A}^k$.

Suppose a solution x^k to

$$\min_{d}\ g^T x + \frac{1}{2}x^T H x$$
$$\text{s.t. } A_i x + b_i = 0, i \in \mathcal{E} \cup \mathcal{A}^k.$$

is feasible, but not optimal.

- Consider $j \in \mathcal{A}^k$ such that $\lambda_j^k < 0$.
- An improving direction is obtained by considering the problem

$$\min_{d}\ g^T(x^k + d) + \frac{1}{2}(x^k + d)^T H (x^k + d)$$
$$\text{s.t. } A_i(x^k + d) + b_i = 0, i \in \mathcal{E} \cup \mathcal{A}^k \setminus \{j\}.$$

or equivalently,

$$\min_{d}\ (g + Hx^k)^T d + \frac{1}{2}d^T H d$$
$$\text{s.t. } A_i d = 0, i \in \mathcal{E} \cup \mathcal{A}^k \setminus \{j\}.$$

- If this problem is unbounded or has a solution $d \neq 0$, then such a d with

$$(g + Hx^k)^T d < 0 \text{ and } A_i d = 0, i \in \mathcal{E} \cup \mathcal{A}^k \setminus \{j\}$$

is an improving direction from x^k.

Do we know if such a direction will maintain feasibility for all constraints?

- If we are not at a degenerate point in that

$$A_i x + b_i < 0, i \in \mathcal{I} \setminus \mathcal{A}_k,$$

  then we maintain feasibility for these constraints for any small displacement.
- If we are at a degenerate point, then we need to worry...
- Finally, what about for the constraint that we are removing from the active set? We do indeed remain feasible; cf. Theorem 16.5 in textbook.

Let $\mathcal{A}_k$ be a guess of the optimal active set corresponding to a feasible x_k for $k = 0, 1, 2, \dots$

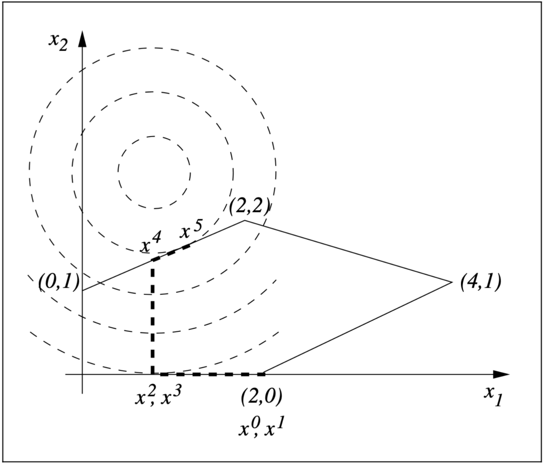
1. Solve the active-set (equality constrained) QOP to obtain (d_k, λ_k)
2. If $d_k = 0$ and $\lambda_k^{\mathcal{I}} > 0$, then stop; x_k is optimal
3. If $d_k \neq 0$ and $x_k + d_k$ is feasible, then set

   $$x_{k+1} \leftarrow x_k + d_k \ \text{ and } \ \mathcal{A}_{k+1} \leftarrow \mathcal{A}_k$$

   and go to step 1.
4. If $d_k \neq 0$ and $x_k + d_k$ is infeasible (for any constraint), then set α_k as the largest value such that $x_{k+1} \leftarrow x_k + \alpha_k d_k$ is feasible, set $\mathcal{A}_k$ as the active set at x_{k+1}, and go to step 1.
5. Choose any j such that $\lambda_k^j < 0$ and set $\mathcal{A}_{k+1} \leftarrow \mathcal{A}_k \setminus \{j\}$
6. Return to step 1.

$$\min_x \; q(x) = (x_1 - 1)^2 + (x_2 - 2.5)^2$$

$$\begin{aligned} \text{subject to} \quad x_1 - 2x_2 + 2 &\geq 0, \\ -x_1 - 2x_2 + 6 &\geq 0, \\ -x_1 + 2x_2 + 2 &\geq 0, \\ x_1 &\geq 0, \\ x_2 &\geq 0. \end{aligned}$$



$$\boldsymbol{x}^{(0)} = (2, 0)^{\mathrm{T}}$$

$$\mathcal{A} = \{3, 5\}$$

1. Example: Markowitz Model
2. Active-set Method
3. Interior Point Method

Consider the convex QP:

$$\begin{aligned} \min\ & c^T x + \frac{1}{2} x^T Q x \\ \text{s.t. } & Ax = b \\ & x \geq 0, \end{aligned}$$

where $c, x \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, $A \in \mathbb{R}^{m \times n}$, $Q \in \mathbb{R}^{n \times n}$.

Assume $Q \succeq 0$. If there exists a feasible solution, then there exists an optimal solution.

Let $y \in \mathbb{R}^m$ and $s \in \mathbb{R}^n_+$ be the multipliers associated with $Ax = b$ and $x \geq 0$, the Lagrangian is

$$L(x, y, s) = c^T x + \frac{1}{2} x^T Q x + y^T (Ax - b) - s^T x.$$

To determine the Lagrangian dual

$$L_D(y,s) = \min_{x \in \mathbb{R}^n} L(x,y,s)$$

we need stationarity with respect to x:

$$\nabla_x L(x,y,s) = c + Qx + A^T y - s = 0 \implies x = -Q^{-1}(c + A^T y - s)$$

The dual problem is then given by

$$\begin{aligned} \max_{y,s} \quad & -\frac{1}{2} y^T (AQ^{-1}A^T) y - b^T y \\ \text{s.t.} \quad & A^T y + s = c \\ & s \geq 0 \end{aligned}$$

Replace the primal QP Consider the convex QP:

$$\begin{aligned} \min\ & c^T x + \frac{1}{2} x^T Q x \\ \text{s.t. } & Ax = b \\ & x \geq 0 \end{aligned}$$

The KKT system:

$$\begin{aligned} c + Qx + A^T y - s = &\ 0 \\ Ax = &\ b \\ XSe = &\ 0 \\ x \geq 0,\ \ s \geq &\ 0 \end{aligned}$$

Apply Newton's method to this.

The perturbed KKT system:

$$\begin{aligned} Qx + A^T y - s + c &= 0 \\ Ax &= b \\ XSe &= \mu e \\ x \geq 0,\ s &\geq 0 \end{aligned}$$

Apply Newton's method to this

$$\begin{bmatrix} Q & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = - \begin{bmatrix} Qx + A^T y - s + c \\ Ax - b \\ XSe - \mu e \end{bmatrix}$$

Obtain a symmetric system by

$$\begin{bmatrix} Q & A^T & I \\ A & 0 & 0 \\ I & 0 & S^{-1}X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = - \begin{bmatrix} Qx + A^T y - s + c \\ Ax - b \\ Se - \mu X^{-1} e \end{bmatrix}$$

Interpretation #2

Replace the primal QP with the barrier QP

$$\begin{aligned} &\min\ c^T x + \frac{1}{2} x^T Q x - \mu \sum_{j=1}^{n} \ln x_j \\ &\text{s.t. } Ax = b \end{aligned}$$

The KKT system of the barrier QP is

$$\begin{aligned} Qx + A^T y + c - \mu X^{-1} e &= 0 \\ Ax &= b \end{aligned}$$

Denote $s = \mu X^{-1} e$, i.e., $XSe = \mu e$.

$$\begin{aligned} Qx + A^T y + c - s &= 0 \\ Ax &= b \\ XSe &= \mu e \end{aligned}$$

Considering the domain of $\ln$, which implies $x \geq 0$ and $s \geq 0$.

Apply Newton's method to this

$$\begin{bmatrix} Q & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = - \begin{bmatrix} Qx + A^T y - s + c \\ Ax - b \\ XSe - \mu e \end{bmatrix}$$

Alternatively, you can also apply to the KKT system of the barrier QP is

$$\begin{aligned} Qx + A^T y + c - \mu X^{-1} e = & \; 0 \\ Ax = & \; b \end{aligned}$$

If apply Newton's method directly,

$$\begin{bmatrix} Q + X^{-2} & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = - \begin{bmatrix} Qx + A^T y + c - \mu X^{-1} e \\ Ax - b \end{bmatrix}$$

1: Initialize $(x_0, y_0, s_0) \in \mathcal{F}^0, \mu_0 = \frac{x_0^T s_0}{n}$, $\alpha_0 = 0.9995$
2: Repeat until optimality

- $k \leftarrow k + 1$
- $\mu_k \leftarrow \sigma\mu_{k-1}$ with $\sigma \in (0, 1)$
- Compute Newton direction $(\Delta x, \Delta y, \Delta s)$.
- Ratio test:

$$\begin{aligned} \alpha_P &= \max\{\alpha > 0 \mid x + \alpha\Delta x \geq 0\} \\ \alpha_D &= \max\{\alpha > 0 \mid s + \alpha\Delta s \geq 0\}. \end{aligned}$$

- Update:

$$\begin{aligned} x_{k+1} &\leftarrow x_k + \alpha_0\alpha_P\Delta x \\ y_{k+1} &\leftarrow y_k + \alpha_0\alpha_D\Delta y \\ s_{k+1} &\leftarrow s_k + \alpha_0\alpha_D\Delta s. \end{aligned}$$