Financial Engineering and Risk Management

Review of linear optimization

Martin Haugh **Garud Iyengar**

Columbia University
Industrial Engineering and Operations Research

Hedging problem

- d assets
- Prices at time $t = 0$: $\mathbf{p} \in \mathbb{R}^d$

- Market in m possible states at time $t = 1$
- Price of asset j in state $i = S_{ij}$

$$\mathbf{S}_j = \begin{bmatrix} S_{1j} \\ S_{2j} \\ \vdots \\ S_{mj} \end{bmatrix} \qquad \mathbf{S} = \begin{bmatrix} \mathbf{S}_1 & \mathbf{S}_2 & \dots & \mathbf{S}_d \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} & \dots & S_{1d} \\ S_{21} & S_{22} & \dots & S_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ S_{m1} & S_{m2} & \dots & S_{md} \end{bmatrix} \in \mathbb{R}^{m \times d}$$

- Hedge an obligation $\mathbf{X} \in \mathbb{R}^m$
  - Have to pay X_i if state i occurs
  - Buy/short sell $\boldsymbol{\theta} = (\theta_1, \dots, \theta_d)^\top$ shares to cover obligation

Hedging problem (contd)

- Position $\boldsymbol{\theta} \in \mathbb{R}^d$ purchased at time $t = 0$
  - $\theta_j =$ number of shares of asset j purchased, $j = 1, \ldots, d$
  - Cost of the position $\boldsymbol{\theta} = \sum_{j=1}^{d} p_j \theta_j = \mathbf{p}^\top \boldsymbol{\theta}$

- Payoff from liquidating position at time $t = 1$
  - payoff y_i in state i: $y_i = \sum_{j=1}^{d} S_{ij} \theta_j$
  - Stacking payoffs for all states: $\mathbf{y} = \mathbf{S}\boldsymbol{\theta}$
  - Viewing the payoff vector $\mathbf{y}$: $\mathbf{y} \in \text{range}(\mathbf{S})$

$$\mathbf{y} = \begin{bmatrix} \mathbf{S}_1 & \mathbf{S}_2 & \ldots & \mathbf{S}_d \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_d \end{bmatrix} = \sum_{j=1}^{d} \theta_j \mathbf{S}_j$$

- Payoff $\mathbf{y}$ hedges $\mathbf{X}$ if $\mathbf{y} \geq \mathbf{X}$.

Hedging problem (contd)

- Optimization problem:

$$\begin{array}{rl} \min & \sum_{j=1}^{d} p_j\theta_j \quad (\equiv \mathbf{p}^\top \boldsymbol{\theta}) \\ \text{subject to} & \sum_{j=1}^{d} S_{ij}\theta_j \geq X_i, \quad i = 1, \ldots, m \quad (\equiv \mathbf{S}\boldsymbol{\theta} \geq \mathbf{X}) \end{array}$$

- Features of this optimization problem
  - Linear objective function: $\mathbf{p}^\top \boldsymbol{\theta}$
  - Linear inequality constraints: $\mathbf{S}\boldsymbol{\theta} \geq \mathbf{X}$

- Example of a linear program
  - Linear objective function: either a min/max
  - Linear inequality and equality constraints

$$\begin{array}{rl} \text{max/min}_{\mathbf{x}} & \mathbf{c}^\top \mathbf{x} \\ \text{subject to} & \mathbf{A}_{eq}\mathbf{x} = \mathbf{b}_{eq} \\ & \mathbf{A}_{in}\mathbf{x} \leq \mathbf{b}_{in} \end{array}$$

Linear programming duality

- Linear program

$$P = \begin{array}{rl} \min_{\mathbf{x}} & \mathbf{c}^\top \mathbf{x} \\ \text{subject to} & \mathbf{A}\mathbf{x} \geq \mathbf{b} \end{array}$$

- Dual linear program

$$D = \begin{array}{rl} \max_{\mathbf{u}} & \mathbf{b}^\top \mathbf{u} \\ \text{subject to} & \mathbf{A}^\top \mathbf{u} = \mathbf{c} \\ & \mathbf{u} \geq \mathbf{0} \end{array}$$

Theorem.

- **Weak Duality**: $P \geq D$
- **Bound**: $\mathbf{x}$ feasible for P, $\mathbf{u}$ feasible for D, $\mathbf{c}^\top \mathbf{x} \geq P \geq D \geq \mathbf{b}^\top \mathbf{u}$
- **Strong Duality**: Suppose P or D finite. Then $P = D$.
- Dual of the dual is the primal (original) problem

More duality results

- Here is another primal-dual pair

$$\begin{array}{rl} \min_{\mathbf{x}} & \mathbf{c}^\top \mathbf{x} \\ \text{subject to} & \mathbf{A}\mathbf{x} = \mathbf{b} \end{array} \quad = \quad \begin{array}{rl} \max_{\mathbf{u}} & \mathbf{b}^\top \mathbf{u} \\ \text{subject to} & \mathbf{A}^\top \mathbf{u} = \mathbf{c} \end{array}$$

- General idea for constructing duals

$$\begin{aligned} P &= \min\{\mathbf{c}^\top \mathbf{x} : \mathbf{A}\mathbf{x} \geq \mathbf{b}\} \\ &\geq \min\{\mathbf{c}^\top \mathbf{x} - \mathbf{u}^\top (\mathbf{A}\mathbf{x} - \mathbf{b}) : \mathbf{A}\mathbf{x} \geq \mathbf{b}\} \text{ for all } \mathbf{u} \geq \mathbf{0} \\ &\geq \mathbf{b}^\top \mathbf{u} + \min\{(\mathbf{c} - \mathbf{A}^\top \mathbf{u})^\top \mathbf{x} : \mathbf{x} \in \mathbb{R}^n\} \\ &= \begin{cases} \mathbf{b}^\top \mathbf{u} & \mathbf{A}^\top \mathbf{u} = \mathbf{c} \\ -\infty & \text{otherwise} \end{cases} \\ &\geq \max\{\mathbf{b}^\top \mathbf{u} : \mathbf{A}^\top \mathbf{u} = \mathbf{c}\} \end{aligned}$$

- Lagrangian relaxation: dualize constraints and relax them!

Financial Engineering and Risk Management

Review of nonlinear optimization

Martin Haugh Garud Iyengar

Columbia University
Industrial Engineering and Operations Research

Unconstrained nonlinear optimization

- Optimization problem

$$\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x})$$

- Categorization of minimum points
  - $\mathbf{x}^*$ global minimum if $f(\mathbf{y}) \geq f(\mathbf{x}^*)$ for all $\mathbf{y}$
  - $\mathbf{x}^*_{loc}$ local minimum if $f(\mathbf{y}) \geq f(\mathbf{x}^*_{loc})$ for all $\mathbf{y}$ such that $\|\mathbf{y} - \mathbf{x}^*_{loc}\| \leq r$

- Sufficient condition for local min
  - gradient $\nabla f(\mathbf{x}) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix} = \mathbf{0}$: local stationarity
  - Hessian $\nabla^2 f(\mathbf{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$ positive semidefinite

- Gradient condition is sufficient if the function $f(\mathbf{x})$ is convex.

Unconstrained nonlinear optimization

- Optimization problem

$$\min_{\mathbf{x}\in\mathbb{R}^2} \; x_1^2 + 3x_1x_2 + x_2^3$$

- Gradient

$$\nabla f(\mathbf{x}) = \begin{bmatrix} 2x_1 + 3x_2 \\ 3x_1 + 3x_2^2 \end{bmatrix} = \mathbf{0} \quad \Rightarrow \quad \mathbf{x} = \mathbf{0}, \quad \begin{bmatrix} -\frac{9}{4} \\ \frac{3}{2} \end{bmatrix}$$

- Hessian at $\mathbf{x}$: $\mathbf{H} = \begin{bmatrix} 2 & 3 \\ 3 & 6x_2 \end{bmatrix}$
  - $\mathbf{x} = \mathbf{0}$: $\mathbf{H} = \begin{bmatrix} 2 & 3 \\ 3 & 0 \end{bmatrix}$. Not positive definite. Not local minimum.
  - $\mathbf{x} = \begin{bmatrix} -\frac{9}{4} \\ \frac{3}{2} \end{bmatrix}$: $\mathbf{H} = \begin{bmatrix} 2 & 3 \\ 3 & 9 \end{bmatrix}$. Positive semidefinite. Local minimum

Lagrangian method

- Constrained optimization problem

$$\begin{aligned} \max_{\mathbf{x}\in\mathbb{R}^2} \quad & 2\ln(1+x_1) + 4\ln(1+x_2), \\ \text{s.t.} \quad & x_1 + x_2 = 12 \end{aligned}$$

- Convex problem. But constraints make the problem hard to solve.
- Form a Lagrangian function

$$\mathcal{L}(\mathbf{x}, v) = 2\ln(1+x_1) + 4\ln(1+x_2) - v(x_1 + x_2 - 12)$$

- Compute the stationary points of the Lagrangian as a function of v

$$\nabla\mathcal{L}(\mathbf{x}, v) = \begin{bmatrix} \frac{2}{1+x_1} - v \\ \frac{4}{1+x_2} - v \end{bmatrix} = \mathbf{0} \quad \Rightarrow \quad x_1 = \frac{2}{v} - 1, \quad x_2 = \frac{4}{v} - 1$$

- Substituting in the constraint $x_1 + x_2 = 12$, we get

$$\frac{6}{v} = 14 \quad \Rightarrow \quad v = \frac{3}{7} \quad \Rightarrow \quad \mathbf{x} = \frac{1}{3}\begin{bmatrix} 11 \\ 25 \end{bmatrix}$$

Portfolio Selection

- Optimization problem

$$\begin{aligned} \max_{\mathbf{x}} \quad & \boldsymbol{\mu}^\top \mathbf{x} - \lambda \mathbf{x}^\top \mathbf{V} \mathbf{x} \\ \text{s.t.} \quad & \mathbf{1}^\top \mathbf{x} = 1 \end{aligned}$$

  Constraints make the problem hard!

- Lagrangian function

$$\mathcal{L}(\mathbf{x}, v) = \boldsymbol{\mu}^\top \mathbf{x} - \lambda \mathbf{x}^\top \mathbf{V} \mathbf{x} - v(\mathbf{1}^\top \mathbf{x} - 1)$$

  - Solve for the maximum value with no constraints

$$\nabla_x \mathcal{L}(\mathbf{x}, v) = \boldsymbol{\mu} - 2\lambda \mathbf{V}\mathbf{x} - v\mathbf{1} = \mathbf{0} \quad \Rightarrow \quad \mathbf{x} = \frac{1}{2\lambda} \cdot \mathbf{V}^{-1}(\boldsymbol{\mu} - v\mathbf{1})$$

  - Solve for v from the constraint

$$\mathbf{1}^\top \mathbf{x} = 1 \quad \Rightarrow \quad \mathbf{1}^\top \mathbf{V}^{-1}(\boldsymbol{\mu} - v\mathbf{1}) = 2\lambda \quad \Rightarrow \quad v = \frac{\mathbf{1}^\top \mathbf{V}^{-1}\boldsymbol{\mu} - 2\lambda}{\mathbf{1}^\top \mathbf{V}^{-1}\mathbf{1}}$$

  - Substitute back in the expression for $\mathbf{x}$