ORF522 – Linear and Nonlinear Optimization

9. Sensitivity analysis for linear optimization

Bartolomeo Stellato — Fall 2022

Ed Forum

- Dual simplex applications?
- In the dual simplex part, we talked about in primal simplex, $X_B > 0$ and $X_N = 0$. However, I can confused about why in dual problem $\bar{c}_B = 0$ and $\bar{c}_N > 0$. Is there any intuition behind this?
- In the illustration of dual simplex method, we used the fact that if y = -A_b^(-T)c_b, then A^Ty + c >=0 is equivalent to reduced cost >=0. From there (page 32), we seem to be constantly using A^Ty + c as the vector of reduced cost. However, I'm wondering why we can use this previous assumption in all our steps during the dual simplex. Why is y = -A_b^(-T)c_b satisfied at all such middle steps or is this something only satisfied at the optimal solution?

Recap

Reduced costs

Interpretation

Change in objective/marginal cost of adding x_j to the basis

$$\bar{c}_j = \boxed{c_j} \boxed{- c_B^T A_B^{-1} A_j}$$

Cost per-unit increase of variable x_j

Cost to change other variables compensating for x_j to enforce $Ax = b$

- $\bar{c}_j > 0$: adding x_j will increase the objective (bad)
- $\bar{c}_j < 0$: adding x_j will decrease the objective (good)

Reduced costs for basic variables is 0

$$\begin{aligned} \bar{c}_{B(i)} &= c_{B(i)} - c_B^T A_B^{-1} A_{B(i)} = c_{B(i)} - c_B^T (A_B^{-1} A_B) e_i \\ &= c_{B(i)} - c_B^T e_i = c_{B(i)} - c_{B(i)} = 0 \end{aligned}$$

Vector of reduced costs

Reduced costs

$$\bar{c}_j = c_j - c_B^T A_B^{-1} A_j$$

Isolate basis B-related components p
(they are the same across j)

$$\bar{c}_j = c_j - A_j^T (A_B^{-1})^T c_B = c_j - A_j^T p$$

Full vector in one shot?

$$\bar{c} = (\bar{c}_1, \dots, \bar{c}_n)$$

Obtain p by solving linear system

$$p = (A_B^{-1})^T c_B \quad \Rightarrow \quad A_B^T p = c_B$$

Note: $(M^{-1})^T = (M^T)^{-1}$
for any square invertible M

Computing reduced cost vector

1. Solve $A_B^T p = c_B$
2. $\bar{c} = c - A^T p$

Primal and dual basic feasible solutions

Primal problem

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{array}$$

Dual problem

$$\begin{array}{ll} \text{maximize} & -b^T y \\ \text{subject to} & A^T y + c \geq 0 \end{array}$$

Given a **basis** matrix A_B

Primal feasible: $Ax = b,\ x \geq 0 \quad \Rightarrow \quad x_B = A_B^{-1} b \geq 0$

Reduced costs

Dual feasible: $A^T y + c \geq 0$. If $y = -A_B^{-T} c_B \quad \Rightarrow \quad c - A^T A_B^{-T} c_B \geq 0$

Zero duality gap: $c^T x + b^T y = c_B^T x_B - b^T A_B^{-T} c_B = c_B^T x_B - c_B^T A_B^{-1} b = 0$

(by construction)

Today's lecture
[Chapter 5, LO]

Sensitivity analysis in linear optimization

- Adding new constraints and variables
- Change problem data
- Differentiable optimization

Adding new constraints and variables

Adding new variables

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{array} \longrightarrow \begin{array}{ll} \text{minimize} & c^T x + c_{n+1} x_{n+1} \\ \text{subject to} & Ax + A_{n+1} x_{n+1} = b \\ & x, x_{n+1} \geq 0 \end{array}$$

Solution $x^\star, y^\star$

Solution $(x^\star, 0), y^\star$ **optimal** for the new problem?

Adding new variables

Optimality conditions

$$\begin{array}{ll} \text{minimize} & c^T x + c_{n+1} x_{n+1} \\ \text{subject to} & Ax + A_{n+1} x_{n+1} = b \\ & x, x_{n+1} \geq 0 \end{array} \longrightarrow$$

Solution $(x^\star, 0)$ is still **primal feasible**

Is $y^\star$ still **dual feasible**?

$$A_{n+1}^T y^\star + c_{n+1} \geq 0$$

Yes

$(x^\star, 0)$ still **optimal** for new problem

Otherwise

Primal simplex

Adding new variables

Example

minimize	$-60x_1 - 30x_2 - 20x_3$	-profit
subject to	$8x_1 + 6x_2 + x_3 \leq 48$	material
	$4x_1 + 2x_2 + 1.5x_3 \leq 20$	production
	$2x_1 + 1.5x_2 + 0.5x_3 \leq 8$	quality control
	$x \geq 0$	

$$c = (-60, -30, -20, 0, 0, 0)$$

minimize	$c^T x$
subject to	$Ax = b$
	$x \geq 0$

$$A = \begin{bmatrix} 8 & 6 & 1 & 1 & 0 & 0 \\ 4 & 2 & 1.5 & 0 & 1 & 0 \\ 2 & 1.5 & 0.5 & 0 & 0 & 1 \end{bmatrix}$$

$$b = (48, 20, 8)$$

$x^\star = (2, 0, 8, 24, 0, 0), \quad y^\star = (0, 10, 10), \quad c^T x^\star = -280,$ basis $\{1, 3, 4\}$

Adding new variables

Example: add new product?

$$c = (-60, -30, -20, 0, 0, 0, -15)$$

minimize	$c^T x + c_{n+1} x_{n+1}$
subject to	$Ax + A_{n+1} x_{n+1} = b$
	$x, x_{n+1} \geq 0$

$$A = \begin{bmatrix} 8 & 6 & 1 & 1 & 0 & 0 & 1 \\ 4 & 2 & 1.5 & 0 & 1 & 0 & 1 \\ 2 & 1.5 & 0.5 & 0 & 0 & 1 & 1 \end{bmatrix}$$

$$b = (48, 20, 8)$$

Previous solution

$x^\star = (2, 0, 8, 24, 0, 0), \quad y^\star = (0, 10, 10), \quad c^T x^\star = -280, \quad$ basis $\{1, 3, 4\}$

Still optimal

$$A_{n+1}^T y^\star + c_{n+1} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 10 \\ 10 \end{bmatrix} - 15 = 5 \geq 0$$

Shall we add a new product?

Adding new constraints

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{array} \longrightarrow \begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & a_{m+1}^T x = b_{m+1} \\ & x \geq 0 \end{array}$$

Solution $x^\star, y^\star$

Dual

$$\begin{array}{ll} \text{maximize} & -b^T y - b_{m+1} y_{m+1} \\ \text{subject to} & A^T y + a_{m+1} y_{m+1} + c \geq 0 \end{array}$$

Solution $x^\star, (y^\star, 0)$ **optimal** for the new problem?

Adding new constraints

Optimality conditions

maximize	$-b^T y - b_{m+1} y_{m+1}$
subject to	$A^T y + a_{m+1} y_{m+1} + c \geq 0$

$\longrightarrow$ Solution $(y^\star, 0)$ is still **dual feasible**

Is $x^\star$ still **primal feasible**?

$$\begin{aligned} &Ax = b \\ &a_{m+1}^T x = b_{m+1} \\ &x \geq 0 \end{aligned}$$

Yes

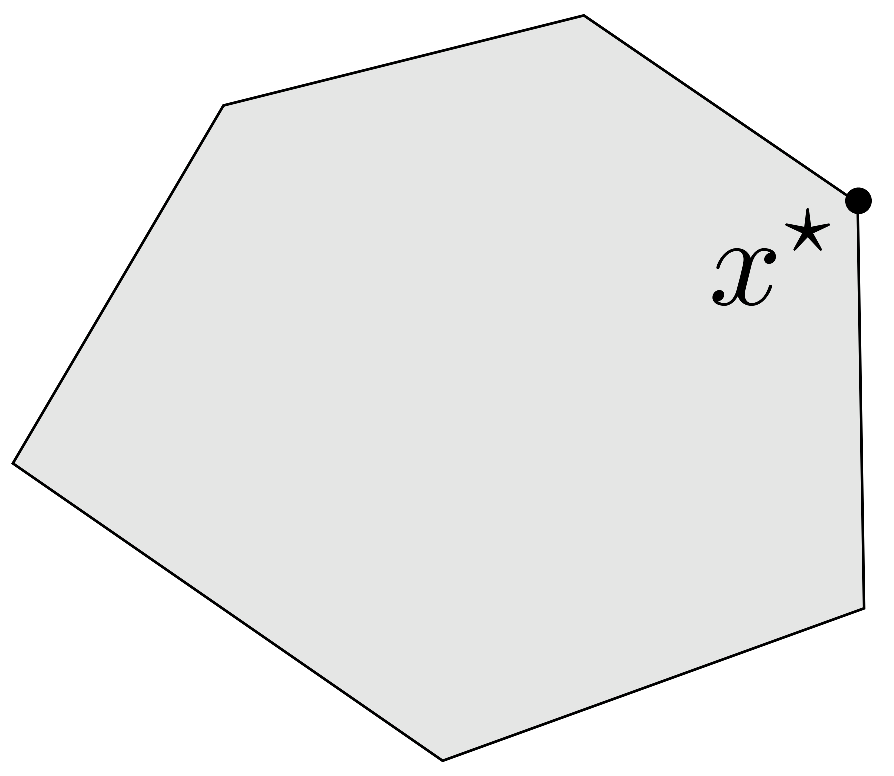
$x^\star$ still **optimal** for new problem

Otherwise

Dual simplex

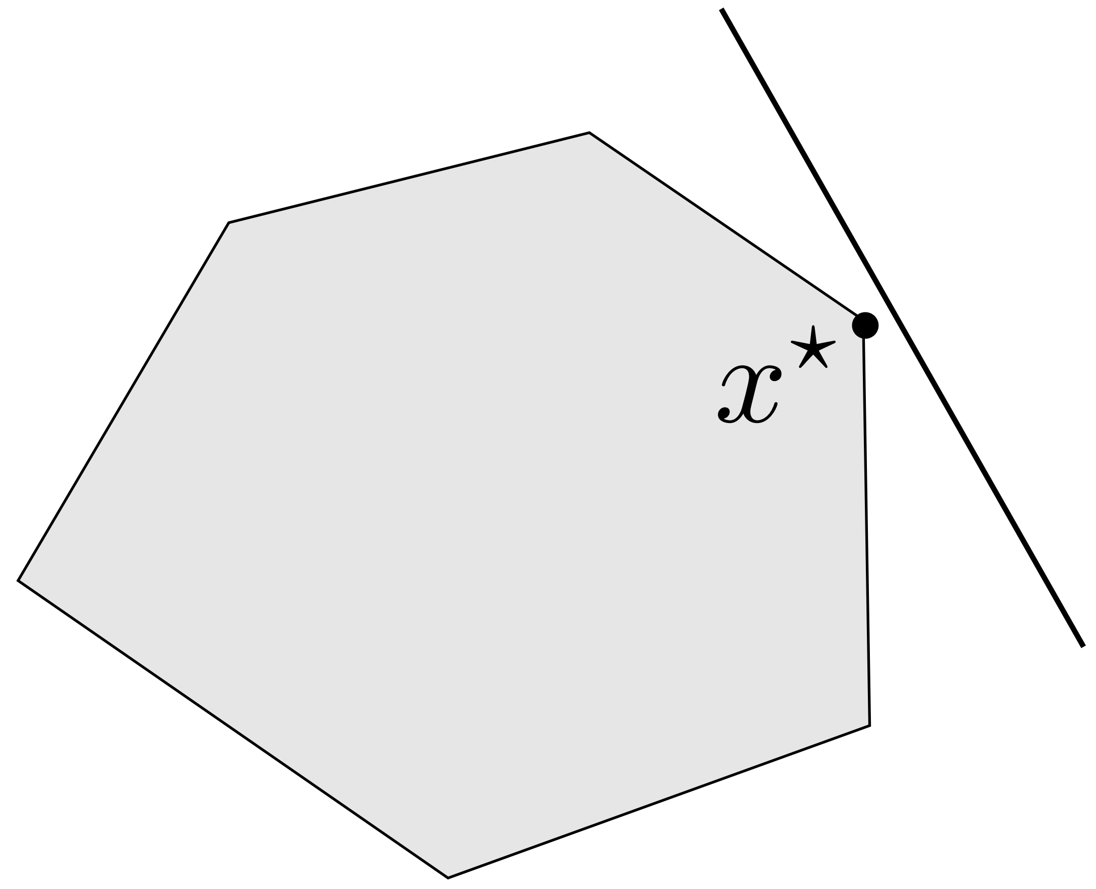
Adding new constraints

Example

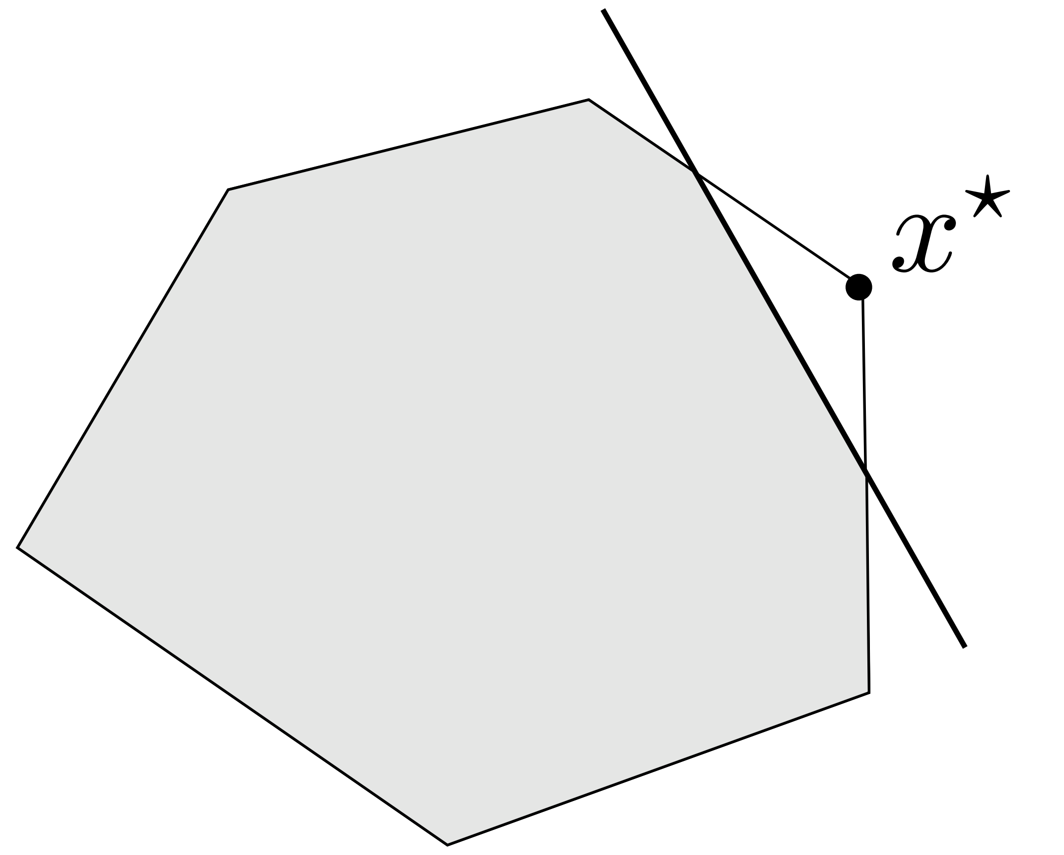


Add new constraint

$x^\star$ still **feasible**



$x^\star$ **infeasible**



Global sensitivity analysis

Information from primal-dual solution

Goal: extract information from $x^\star, y^\star$ about their sensitivity with respect to changes in problem data

Modified LP

$$
\begin{array}{ll}
\text{minimize} & c^T x \\
\text{subject to} & Ax = b + u \\
& x \geq 0
\end{array}
$$

Optimal cost $p^\star(u)$

Global sensitivity

Dual of modified LP

$$\begin{array}{ll} \text{maximize} & -(b+u)^T y \\ \text{subject to} & A^T y + c \geq 0 \end{array}$$

Global lower bound

Given $y^\star$ a dual optimal solution for $u = 0$, then

$$\begin{aligned} p^\star(u) &\geq -(b+u)^T y^\star \\ &= p^\star(0) - u^T y^\star \end{aligned}$$

(from weak duality and dual feasibility)

It holds for any u

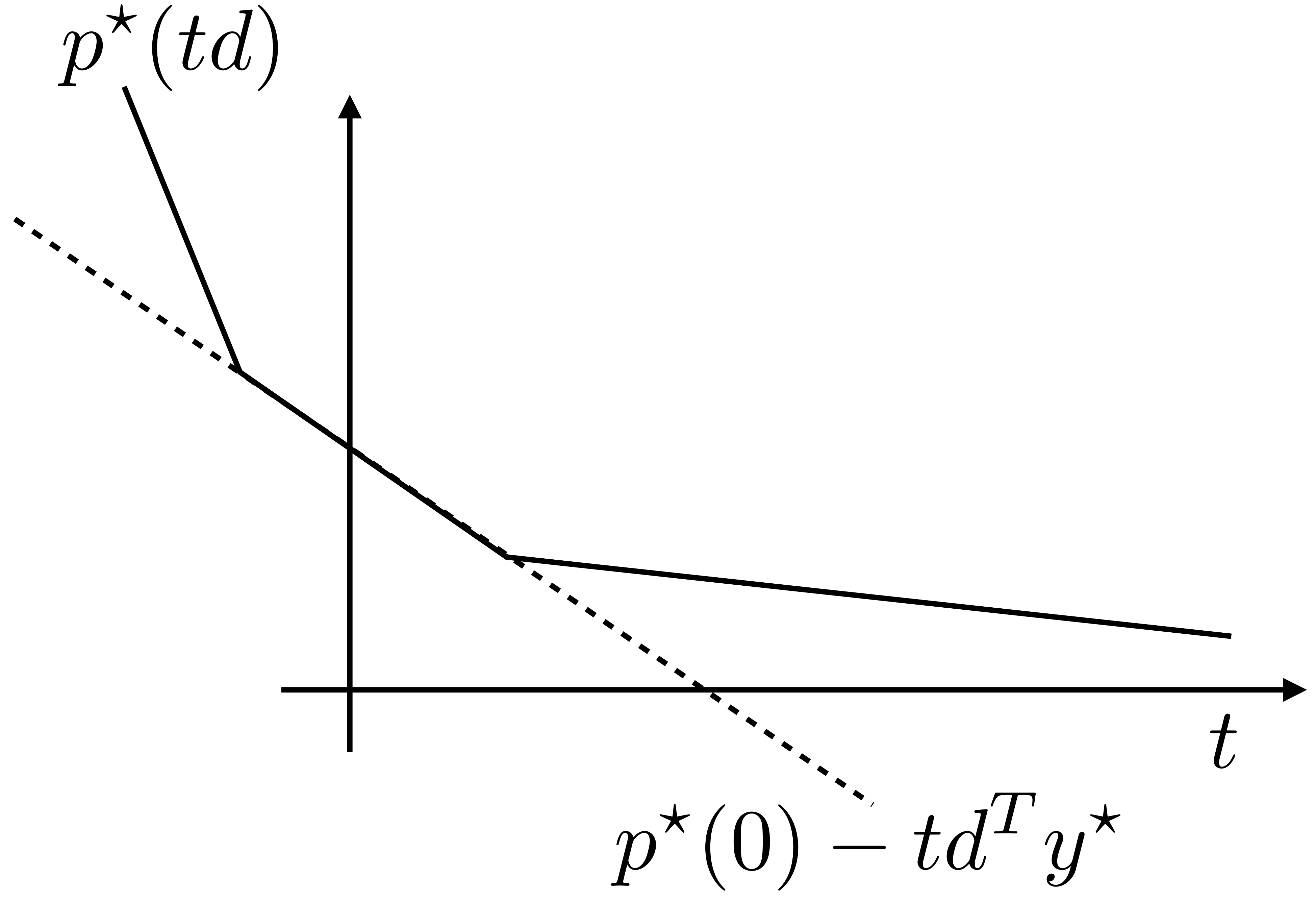
Global sensitivity

Example

Take $u = td$ with $d \in \mathbf{R}^m$ fixed

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b + td \\ & x \geq 0 \end{array}$$

$p^\star(td)$ is the optimal value as a function of t



Sensitivity information (assuming $d^T y^\star \geq 0$)

- $t < 0$ the optimal value increases
- $t > 0$ the optimal value decreases (not so much if t is small)

Optimal value function

$$p^\star(u) = \min\{c^T x \mid Ax = b + u,\ x \geq 0\}$$

Assumption: $p^\star(0)$ is finite

Properties

- $p^\star(u) > -\infty$ everywhere (from global lower bound)
- the domain $\{u \mid p^\star(u) < +\infty\}$ is a polyhedron
- $p^\star(u)$ is piecewise-linear on its domain

Optimal value function is piecewise linear

Proof

Dual feasible set

$$p^\star(u) = \min\{c^T x \mid Ax = b + u,\ x \geq 0\}$$

$$D = \{y \mid A^T y + c \geq 0\}$$

Assumption: $p^\star(0)$ is finite

If $p^\star(u)$ finite

$$p^\star(u) = \max_{y \in D} -(b+u)^T y = \max_{k=1,\dots,r} -y_k^T u - b^T y_k$$

$y_1, \dots, y_r$ are the extreme points of D

Local sensitivity analysis

Local sensitivity

u in neighborhood of the origin

Original LP

minimize $\quad c^T x$

subject to $\quad Ax = b$

$\quad x \geq 0$

$\longrightarrow$

Optimal solution

Primal $\quad x_i = 0, \quad i \notin B$

$x_B^\star = A_B^{-1} b$

Dual $\quad y^\star = -A_B^{-T} c_B$

Modified LP

minimize $\quad c^T x$

subject to $\quad Ax = b + u$

$\quad x \geq 0$

Modified dual

maximize $\quad -(b+u)^T y$

subject to $\quad A^T y + c \geq 0$

Optimal basis does not change

Modified optimal solution

$$x_B^\star(u) = A_B^{-1}(b+u) = x_B^\star + A_B^{-1} u$$

$$y^\star(u) = y^\star$$

Derivative of the optimal value function

Modified optimal solution

$$x_B^\star(u) = A_B^{-1}(b+u) = x_B^\star + A_B^{-1}u$$
$$y^\star(u) = y^\star$$

Optimal value function

$$\begin{aligned} p^\star(u) &= c^T x^\star(u) \\ &= c^T x^\star + c_B^T A_B^{-1} u \\ &= p^\star(0) - y^{\star T} u \end{aligned}$$

(affine for small u)

Local derivative

$$\frac{\partial p^\star(u)}{\partial u} = -y^\star$$

($y^\star$ are the **shadow prices**)

Sensitivity example

minimize	$-60x_1 - 30x_2 - 20x_3$	-profit
subject to	$8x_1 + 6x_2 + x_3 \leq 48$	material
	$4x_1 + 2x_2 + 1.5x_3 \leq 20$	production
	$2x_1 + 1.5x_2 + 0.5x_3 \leq 8$	quality control
	$x \geq 0$	

$x^\star = (2, 0, 8, 24, 0, 0), \quad y^\star = (0, 10, 10), \quad c^T x^\star = -280,$ basis $\{1, 3, 4\}$

What does $y_3^\star = 10$ mean?

Let's increase the quality control budget by 1, i.e., $u = (0, 0, 1)$

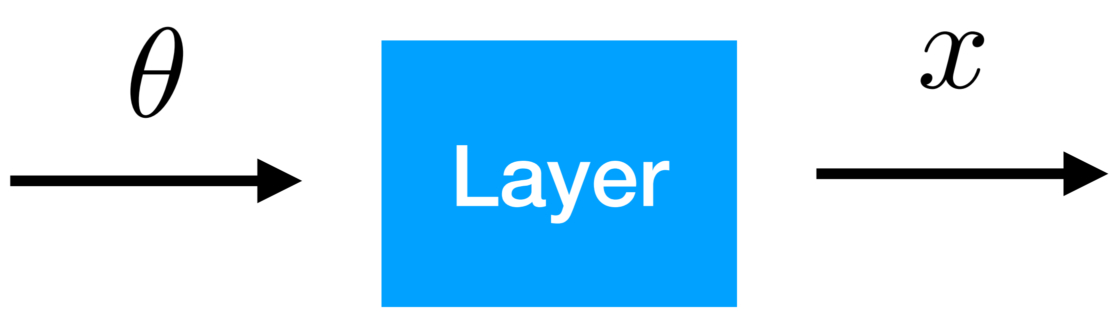
$$p^\star(10) = p^\star(0) - y^{\star T} u = -280 - 10 = -290$$

Differentiable optimization

Training a neural network

Single layer model

$$x = f(\theta)$$



Training

minimize $\quad \mathcal{L}(\theta)$

Gradient descent (more on this later)

$$\theta \leftarrow \theta - t\nabla_\theta \mathcal{L}(\theta)$$

Sensitivity

$$\nabla_\theta \mathcal{L} = \left(\frac{\partial \mathcal{L}}{\partial \theta}\right)^T = \left(\frac{\partial \mathcal{L}}{\partial x}\frac{\partial x}{\partial \theta}\right)^T = \left(\frac{\partial x}{\partial \theta}\right)^T \nabla_x \mathcal{L}$$

Can f be an **optimization problem**?

Implicit layers

https://implicit-layers-tutorial.org/

find $\quad x(\theta)$

subject to $\quad r(\theta, x(\theta)) = 0$

($x(\theta)$ is implicitly defined by r)

How do we compute derivatives?

$$\frac{\partial x(\theta)}{\partial \theta}$$

Implicit function theorem

Under mild assumptions (non-singularity),

$$\frac{\partial r(\theta, x(\theta))}{\partial x} \frac{\partial x(\theta)}{\partial \theta} + \frac{\partial r(\theta, x(\theta))}{\partial \theta} = 0 \longrightarrow \frac{\partial x(\theta)}{\partial \theta} = -\left(\frac{\partial r(\theta, x(\theta))}{\partial x}\right)^{-1} \frac{\partial r(\theta, x(\theta))}{\partial \theta}$$

[Theorem 1. B.1, Dontchev and Rockafellar 2009]

Optimization layers

$$x^\star(\theta) = \begin{array}{ll} \underset{x}{\text{argmin}} & c^T x \\ \text{subject to} & Ax \le b \end{array}$$

Parameters: $\theta = \{c, A, b\}$
Solution $x^\star(\theta)$

Features

- Add **domain knowledge** and **hard constraints**
- **End-to-end** training and optimization
- Nice theory and algorithms for general **convex optimization**
- **Applications** in RL, control, meta-learning, game theory, etc.

Goal

Compute $\dfrac{\partial x^\star(\theta)}{\partial \theta}$

Optimality conditions

minimize $\quad c^T x$
subject to $\quad Ax \leq b$

Parameters: $\theta = \{c, A, b\}$
Solution $x^\star(\theta)$

Solve and obtain primal-dual pair $x^\star, y^\star$ (forward-pass)

Optimality conditions

$$\begin{aligned} A^T y + c = 0 \\ \mathbf{diag}(y)(Ax - b) = 0 \\ y \geq 0,\ b - Ax \geq 0 \end{aligned}$$

Mapping $r(\theta, x(\theta)) = 0$

Computing derivatives

Take differentials

$$\begin{aligned} A^T y^\star + c &= 0 \\ \mathbf{diag}(y^\star)(Ax - b) &= 0 \end{aligned} \longrightarrow \begin{aligned} \mathrm{d}A^T y^\star + A^T \mathrm{d}y &= 0 \\ \mathbf{diag}(Ax - b)\mathrm{d}y + \mathbf{diag}(y^\star)(\mathrm{d}Ax^\star + A\mathrm{d}x - \mathrm{d}b) + \mathrm{d}c &= 0 \end{aligned}$$

Linear system

$$\begin{bmatrix} 0 & A^T \\ \mathbf{diag}(y^\star)A & \mathbf{diag}(Ax^\star - b) \end{bmatrix} \begin{bmatrix} \mathrm{d}x \\ \mathrm{d}y \end{bmatrix} = - \begin{bmatrix} \mathrm{d}A^T y^\star + \mathrm{d}c \\ \mathbf{diag}(y^\star)(\mathrm{d}Ax^\star - \mathrm{d}b) \end{bmatrix}$$

Example: How does $x^\star$ change with b_1?

Set $\mathrm{d}b = e_1, \mathrm{d}A = 0, \mathrm{d}c = 0$ and solve the linear system.

The solution $\mathrm{d}x$ will correspond to $\dfrac{\partial x}{\partial b_1}$

Is it always differentiable?

The linear system matrix must be invertible
(the problem must have unique solution)

$$\underbrace{\begin{bmatrix} 0 & A^T \\ \mathbf{diag}(y^\star)A & \mathbf{diag}(Ax^\star - b) \end{bmatrix}}_{M} \begin{bmatrix} \mathrm{d}x \\ \mathrm{d}y \end{bmatrix} = -\underbrace{\begin{bmatrix} \mathrm{d}A^T y^\star + \mathrm{d}c \\ \mathbf{diag}(y^\star)(\mathrm{d}Ax^\star - \mathrm{d}b) \end{bmatrix}}_{q}$$

Remember. implicit function theorem

$$\frac{\partial x(\theta)}{\partial \theta} = -\left(\frac{\partial r(\theta, x(\theta))}{\partial x}\right)^{-1} \frac{\partial r(\theta, x(\theta))}{\partial \theta}$$

If not, **least squares "**subdifferential"

$$\text{minimize} \quad \left\| M \begin{bmatrix} \mathrm{d}x \\ \mathrm{d}y \end{bmatrix} + q \right\|_2^2$$

Example

Learning to play Sudoku

			3
1			
		4	
4			1

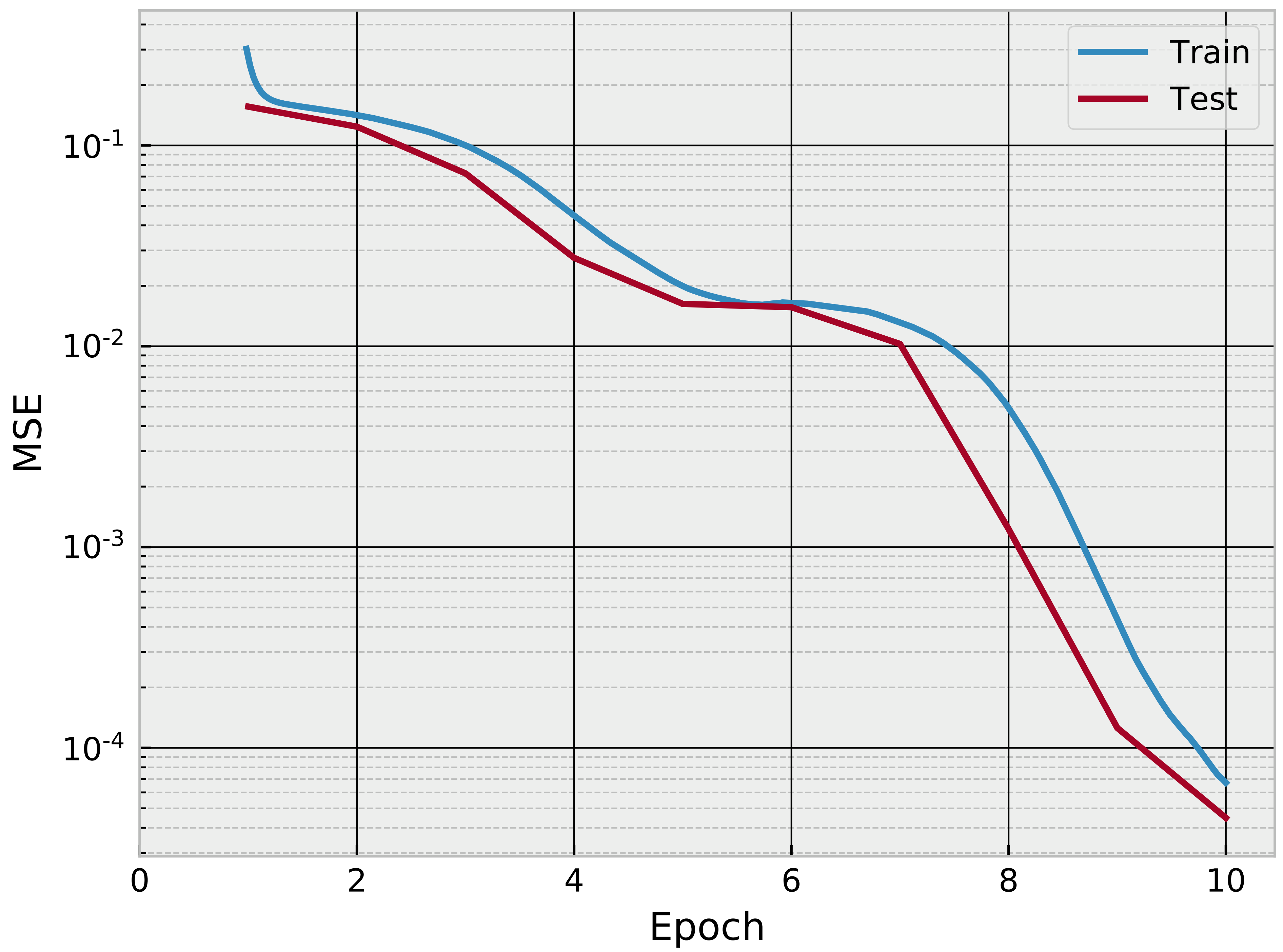
2	4	1	3
1	3	2	4
3	1	4	2
4	2	3	1

Sudoku constraint satisfaction problem

$$\begin{array}{ll} \text{minimize} & 0 \\ \text{subject to} & Ax = b \\ & x \geq 0, \;\; x \in \mathbf{Z}^d \end{array}$$

Linear optimization layer (parameters $\theta = \{A, b\}$)

$$x^\star = \begin{array}{ll} \underset{x}{\text{argmin}} & 0 \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{array}$$



[OptNet: Differentiable Optimization as a Layer in Neural Networks, B. Amos and J. Z. Kolter ICML 2017]

Sensitivity analysis in linear optimization

Today, we learned to:

- **Use** the most appropriate primal/dual simplex algorithm when variables and/or constraints are added

- **Analyze** sensitivity of the cost with respect to change in the data

- **Apply** sensitivity analysis to differentiable linear optimization layers

Next lecture

- Barrier methods for linear optimization