ORF 522

Linear Programming and Convex Analysis

Duality

Marco Cuturi

Today

- Duality theory, the general case.
  - Lagrangian
  - Lagrange dual function and optima for the dual.
  - Weak and strong duality.
- A closer look at the linear case.
  - dual programs for LP's
  - Weak duality and two corollaries
  - Strong Duality
  - Complementary Slackness
- Examples
  - Simple LP.
  - Max-flow / min-cut problem.

Duality

- **Duality theory**:
  - Keep this in mind: only a long list of **simple** inequalities. . . .
  - In the end: very powerful results at low technical/numerical cost.
  - A few important, intuitive theorems.
  - We provide proofs for LP's here, some more advanced results exist in convexity.

- **In a LP context**:
  - Dual problem provides a different **interpretation** on the same problem.
  - Essentially assigns cost ("displeasure" measure) to constraints.
  - Allow us to study cheaply the sensitivity of the solution to changes in constraints.
  - Provides alternative algorithms (dual-simplex).

Duality : the general case

Optimization problem

- Consider the following **mathematical program**:

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \le 0, \quad i = 1, \dots, m \\ & h_i(x) = 0, \quad i = 1, \dots, p \end{array}$$

  where $x \in \mathcal{D} \subset \mathbf{R}^n$ with optimal value $p^\star$.

- No particular assumptions on $\mathcal{D}$ and the functions f and h (convexity, linearity, continuity, etc)
- Very generic (includes linear programming and many other problems)

Lagrangian

We form the **Lagrangian** of this problem:

$$L(\mathbf{x}, \lambda, \mu) = f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}) + \sum_{i=1}^{p} \mu_i h_i(\mathbf{x})$$

as a function of the original variable $\mathbf{x} \in \mathbf{R}^n$, and additional variables $\lambda \in \mathbf{R}^m$ and $\mu \in \mathbf{R}^p$, called **Lagrange multipliers**.

- The Lagrangian is a **penalized** version of the original objective
- The Lagrange multipliers λ_i, μ_i control the weight of the penalty assigned to each violation.
- The Lagrangian is a smoothed version of the hard problem, we have turned $\mathbf{x} \in C$ into penalties that take into account the constraints that **define** C.
- The idea of replacing **hard** constraints by **penalizations** or **soft** constraints will come again when we will study IPM.

Lagrange dual function

- We originally have

$$L(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f_0(\boldsymbol{x}) + \sum_{i=1}^{m} \boldsymbol{\lambda_i} f_i(\boldsymbol{x}) + \sum_{i=1}^{p} \boldsymbol{\mu_i} h_i(\boldsymbol{x})$$

- The penalized problem is here:

$$\begin{aligned} g(\boldsymbol{\lambda}, \boldsymbol{\mu}) &= \inf_{\boldsymbol{x} \in \mathcal{D}} L(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) \\ &= \inf_{x \in \mathcal{D}} f_0(x) + \textstyle\sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \mu_i h_i(x) \end{aligned}$$

- The function $g(\lambda, \mu)$ is called the **Lagrange dual function**.
  - Easier to solve than the original one (the constraints are gone)
  - Can often be computed explicitly (more later)

Lower bound

- The function $g(\lambda, \mu)$ produces a lower bound on $p^\star$.
- **Lower bound property**: If $\lambda \geq 0$, then $g(\lambda, \mu) \leq p^\star$
    - Why? If $\tilde{x}$ is feasible,
        - ▷ $f_i(\tilde{x}) \leq 0$ and thus $\lambda_i f_i(\tilde{x}) \leq 0$
        - ▷ $h_i(x) = 0$, and thus $\mu_i h_i(\tilde{x}) = 0$
    - thus by construction of L:

$$g(\lambda, \mu) = \inf_{x \in \mathcal{D}} L(x, \lambda, \mu) \leq L(\tilde{x}, \lambda, \mu) \leq f_0(\tilde{x})$$

    - This is true for any feasible $\tilde{x}$, so it must be true for the optimal one, which means $g(\lambda, \mu) \leq f_0(x^\star) = p^\star$.

Lower bound

- We now have a **systematic** way of producing **lower bounds** on the optimal value $p^\star$ of the original problem:

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \le 0, \quad i = 1, \dots, m \\ & h_i(x) = 0, \quad i = 1, \dots, p \end{array}$$

- All it takes is a **feasible point** $\tilde{x} \in \mathcal{D}$, which satisfies:

$$\begin{array}{l} f_i(\tilde{x}) \le 0, \quad i = 1, \dots, m \\ h_i(\tilde{x}) = 0, \quad i = 1, \dots, p \end{array}$$

- We can look for the best possible one. . .

Dual problem

- We can define the **Lagrange dual** problem:

$$\begin{array}{ll} \text{maximize} & g(\lambda, \mu) \\ \text{subject to} & \lambda \geq 0 \end{array}$$

  in the variables $\lambda \in \mathbf{R}^m$ and $\mu \in \mathbf{R}^p$.

- Finds the best, that is **highest**, possible lower bound $g(\lambda, \mu)$ on the optimal value $p^\star$ of the original (now called **primal**) problem.

- We call its optimal value $d^\star$

Dual problem

- For each given x, the function

$$L(x, \lambda, \mu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \mu_i h_i(x)$$

  is **linear** in the variables λ and μ.

- This means that the function

$$g(\lambda, \mu) = \inf_{x \in \mathcal{D}} L(x, \lambda, \mu)$$

  is a minimum of linear functions of (λ, μ), so it must be **concave** in (λ, μ)

- This means that the dual problem is always a **concave maximization** problem, whatever f, g, h's properties are.

Weak duality

We have shown the following property called **weak duality**:

$$d^\star \leq p^\star$$

i.e. the optimal value of the dual is always less than the optimal value of the primal problem.

- We haven't made any further assumptions on the problem
- Weak duality must **always hold**
- Produces lower bounds on the problem at low cost

What happens when $d^\star = p^\star$?. . .

Strong duality

When $d^\star = p^\star$ we have **strong duality**.

- Because $d^\star$ is a lower bound on the optimal value $p^\star$, if both are equal for some (x, λ, μ), the current point must be optimal
- The converse is false: (x, λ, μ) could be optimal with $d^\star < p^\star$
- For most convex problems, we have strong duality
- The difference $p^\star - d^\star$ is called the **duality gap** and is a measure of how optimal the current solution (x, λ, μ).

Slater's conditions

Example of sufficient conditions for **strong duality**:

- **Slater's conditions**. Consider the following problem:

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & Ax = \mathbf{b}, \quad i = 1, \ldots, p \end{array}$$

  where all the $f_i(x)$ are **convex** and assume that:

$$\text{there exists } x \in \mathcal{D}: \; f_i(x) < 0, \; Ax = \mathbf{b}, \quad i = 1, \ldots, m$$

  in other words there is a **strictly feasible point**, then strong duality holds.

- Many other versions exist. . .
- Often easy to check.
- Let's see for linear programs.

Duality: linear programming

Duality: linear programming

- Take a **linear program** in standard form:

$$\begin{array}{ll} \text{minimize} & \mathbf{c}^T\mathbf{x} \\ \text{subject to} & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq 0 (-\mathbf{x} \leq 0) \end{array}$$

- We can form the **Lagrangian**:

$$L(\mathbf{x}, \lambda, \mu) = \mathbf{c}^T\mathbf{x} - \lambda^T\mathbf{x} + \mu^T(A\mathbf{x} - \mathbf{b})$$

- and the **Lagrange dual function**:

$$\begin{aligned} g(\lambda, \mu) &= \inf_{\mathbf{x}} L(\mathbf{x}, \lambda, \mu) \\ &= \inf_{\mathbf{x}} \mathbf{c}^T\mathbf{x} - \lambda^T\mathbf{x} + \mu^T(A\mathbf{x} - b) \end{aligned}$$

Duality: linear programming

- For linear programs, the Lagrange dual function can be computed **explicitly**:

$$
\begin{aligned}
g(\lambda, \mu) &= \inf_{\mathbf{x}} \mathbf{c}^T\mathbf{x} \; - \lambda^T\mathbf{x} + \mu^T(A\mathbf{x} - b) \\
&= \inf_{\mathbf{x}} (c - \lambda + A^T\mu)^T\mathbf{x} - \mathbf{b}^T\mu
\end{aligned}
$$

- This is either $-\mathbf{b}^T\mu$ or $-\infty$, so we finally get:

$$
g(\lambda, \mu) = \begin{cases} -\mathbf{b}^T\mu & \text{if } c - \lambda + A^T\mu = 0 \\ -\infty & \text{otherwise} \end{cases}
$$

- If $g(\lambda, \mu) = -\infty$ we say that (λ, μ) are outside the domain of the dual.

Duality: linear programming

- With $g(\lambda, \mu)$ given by:

$$g(\lambda, \mu) = \begin{cases} -\mathbf{b}^T \mu & \text{if } c - \lambda + A^T \mu = 0 \\ -\infty & \text{otherwise} \end{cases}$$

- we can write the dual program as:

$$\begin{array}{ll} \text{maximize} & g(\lambda, \mu) \\ \text{subject to} & \lambda \geq 0 \end{array}$$

- which is again, writing the domain explicitly:

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

Duality: linear programming

- After simplification:

$$\begin{cases} c - \lambda + A^T\mu = 0 \\ \lambda \geq 0 \end{cases} \iff c + A^T\mu \geq 0$$

- we conclude that the dual of the linear program:

$$\begin{array}{ll} \text{minimize} & \mathbf{c}^T\mathbf{x} \\ \text{subject to} & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq 0 \end{array} \qquad \textbf{(primal)}$$

- is given by:

$$\begin{array}{ll} \text{maximize} & -\mathbf{b}^T\mu \\ \text{subject to} & -A^T\mu \leq c \end{array} \qquad \textbf{(dual)}$$

- equivalently:

$$\begin{array}{ll} \text{maximize} & \mathbf{b}^T\mu \\ \text{subject to} & A^T\mu \leq c \end{array}$$

Dual Linear Program

Up to now, what have we introduced?

- A vector of parameters $\mu \in \mathbf{R}^m$, **one coordinate by constraint**.
- For **any** μ and any feasible $\mathbf{x}$ of the primal $=$ a lower bound on the primal.
- For **some** μ the lower bound is $-\infty$, not useful.
- The **dual problem** computes the **biggest** lower bound.
- We discard values of μ which give $-\infty$ lower bounds.
- This the way **dual constraints** are defined.
- The **dual** is **another linear program** in dimensions $\mathbf{R}^{n\times m}$, that is n constraints and m variables.

From Primal to Dual for general LP's

- Some notations: for $A \in \mathbf{R}^{m \times n}$ we write
  - $\mathbf{a}_j$ for the n column vectors
  - $\mathbf{A}_i$ for the m row vectors of A.
- Following a similar reasoning we can flip from primal to dual changing
  - the constraints linear relationships A,
  - the constraints constants $\mathbf{b}$,
  - the constraints directions $(\leq, \geq, =)$
  - non-negativity conditions,
  - the objective

minimize	$\mathbf{c}^T\mathbf{x}$		maximize	$\mu^T\mathbf{b}$	
subject to	$\mathbf{A}_i^T\mathbf{x} \geq b_i,$	$i \in M_1$	subject to	$\mu_i \geq 0$	$i \in M_1$
	$\mathbf{A}_i^T\mathbf{x} \leq b_i,$	$i \in M_2$		$\mu_i \leq 0$	$i \in M_2$
	$\mathbf{A}_i^T\mathbf{x} = b_i,$	$i \in M_3$		μ_i free	$i \in M_3$
	$x_j \geq 0$	$j \in N_1$		$\mu^T\mathbf{a}_j \leq c_j$	$j \in N_1$
	$x_j \leq 0$	$j \in N_1$		$\mu^T\mathbf{a}_j \geq c_j$	$j \in N_2$
	x_j free	$j \in N_1$		$\mu^T\mathbf{a}_j = c_j$	$j \in N_3$

(1)

Dual Linear Program

- In summary, for any kind of constraint,

primal	minimize	maximize	**dual**
constraints	$\geq b_i$ $\leq b_i$ $= b_i$	≥ 0 ≤ 0 free	variables
variables	≥ 0 ≤ 0 free	$\leq c_j$ $\geq c_j$ $= c_j$	constraints

- For simple cases and in matrix form,

minimize $\mathbf{c}^T\mathbf{x}$ subject to $A\mathbf{x} = \mathbf{b}$ $\mathbf{x} \geq 0$	$\Rightarrow$	maximize $\mathbf{b}^T\mu$ subject to $A^T\mu \leq c$
minimize $\mathbf{c}^T\mathbf{x}$ subject to $A\mathbf{x} \geq \mathbf{b}$	$\Rightarrow$	maximize $\mathbf{b}^T\mu$ subject to $A^T\mu = c$ $\mu \geq 0$

Dual Linear Program: Equivalence Theorems

Theorem 1. *If we transform the dual problem into an equivalent minimization problem and the form its dual, we obtain a problem that is equivalent to the original problem*

- The **dual of the dual** of a given primal LP **is the primal LP** itself.
- Linear programs are **self-dual**.
- Not true in the general case. The dual of the dual is called the **bi-dual** problem.
- The tables before can be used in both directions indifferently.

Dual Linear Program: Equivalence Theorems

Theorem 2. *If we transform a LP (1) into another LP (2) through any of the following operations:*

- *replace free variables with the difference of two nonnegative variables;*
- *replace inequality constraints by an equality constraint with a surplus/slack variable;*
- *remove redundant (colinear) rows of the constraint matrix for standard forms;*

then the duals of (1) and (2) are equivalent, i.e. they are either both infeasible or have the same optimal objective.

Duality for LP's : Weak Duality

We proved weak duality for general programs. Although LP's are a **particular case** the arguments are here explicit:

Theorem 3. *If* $\mathbf{x}$ *is a feasible solution to a primal LP and* μ *is a feasible solution to the dual problem then*

$$\mu^T \mathbf{b} \leq \mathbf{c}^T \mathbf{x}$$

- **Proof idea** check what is called the complementary slackness variables $\mu_i(\mathbf{A}_i^T \mathbf{x} - b_i)$ and $(c_j - \mu^T \mathbf{a}_j)\mathbf{x}_j$ and use the primal/dual relationships.

Weak Duality Proof

Proof. • Let $\mathbf{x} \in \mathbf{R}^n$ and $\mu \in \mathbf{R}^m$ and define

$$\begin{aligned} u_i &= \mu_i(\mathbf{A}_i^T \mathbf{x} - b_i) \quad & i = 1, .., m \\ v_j &= (c_j - \mu^T \mathbf{a}_j)\mathbf{x}_j \quad & j = 1, .., n \end{aligned}$$

- Suppose $\mathbf{x}$ and μ are primal and dual feasible for an LP involving A, $\mathbf{b}$ and $\mathbf{c}$.
- Check Equations 1. Whatever the constraints are,
  - μ_i and $(\mathbf{A}_i^T \mathbf{x} - b_i)$ have the same sign or their product is zero.
  - The same goes for $(c_j - \mu^T \mathbf{a}_j)$ and $\mathbf{x}_j$.
- Hence $u_i, v_j \geq 0$.
- Furthermore $\sum_i^m u_i = \mu^T(A\mathbf{x} - \mathbf{b})$ and $\sum_j^n v_j = (\mathbf{c}^T - \mu^T A)\mathbf{x}$
- Hence $0 \leq \sum_i^m u_i + \sum_j^n v_j = \mathbf{c}^T \mathbf{x} - \mu^T \mathbf{b}$

■

Weak Duality

- Not a very strong result at first look.
- Specially since we already discussed **strong duality**...

- Yet weak duality provides us with the two simple yet **important corollaries**.
- In the following we assume that the **primal** is a **minimization**.
- As usual, results can be easily proved the other way round.

Weak Duality Corollary 1

Corollary 1.
- *If the objective in the primal can be arbitrarily small then the dual problem must be infeasible.*
- *If the objective in the primal can be arbitrarily big then the dual problem must be infeasible.*

Proof.
- By weak duality, $\mu^T \mathbf{b} \leq \mathbf{c}^T \mathbf{x}$ for any two feasible points $\mathbf{x}, \mu$.
- If the objective for feasible $\mathbf{x}$ can be set arbitrarily low, then a feasible μ cannot exist.
- The same applies for a feasible $\mathbf{x}$ if the dual objective can be arbitrarily high.

■

Weak Duality Corollary 2

Corollary 2. *Let* $\mathbf{x}^\star$ *and* $\mu^\star$ *be two feasible solutions to the primal and dual respectively. Suppose that* $\mu^{\star T}\mathbf{b} = \mathbf{c}^T\mathbf{x}^\star$*. Then* $\mathbf{x}^\star$ *and* $\mu^\star$ *are optimal solutions for the primal and dual respectively.*

Proof. For every feasible point of the primal $\mathbf{y}$, $\mathbf{c}^T\mathbf{x}^\star = \mu^{\star T}\mathbf{b} \leq \mathbf{c}^T\mathbf{y}$ hence $\mathbf{x}^\star$ is optimal. Same thing for $\mu^\star$. ■

- Let's check whether strong duality holds or not for linear programs...

Strong Duality

- For linear programs, **strong duality is always ensured**.
- We use the **simplex**'s convergence to the optimal solution in this proof.
- We will cover a more geometric approach in the next lecture.

Theorem 4. *if an LP has an optima, so does its dual, and their* ***respective optimal objectives are equal****.*

- **Proof strategy**:
  - prove it first for a **standard form LP**, showing that the **reduced cost coefficient** can be used to define a **dual feasible solution**..
  - For a general LP, use Theorem 2

Strong Duality: Proof 1

Proof. • Consider the standard form

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

- Let's use the simplex with the lexicographic rule for instance. Let $\mathbf{x}$ be the optimal solution with basis $\mathbf{I}$ and objective z.
- The reduced costs must be nonnegative (here we have a **min** problem) hence

$$\mathbf{c}^T - \mathbf{c}_\mathbf{I}^T B_\mathbf{I}^{-1} A \geq \mathbf{0}^T$$

- Let $\mu^T = \mathbf{c}_\mathbf{I}^T B_\mathbf{I}^{-1}$. Then $\mu^T A \geq \mathbf{c}^T$ coordinate wise.
- μ is a **feasible** solution to the dual problem.
- Furthermore $\mu^T \mathbf{b} = \mathbf{c}_\mathbf{I}^T B_\mathbf{I}^{-1} \mathbf{b} = \mathbf{c}_\mathbf{I}^T \mathbf{x}_\mathbf{I} = z$.
- μ is thus optimal w.r.t to the dual following the previous corollary.

Strong Duality: Proof 2

- Suppose now that we have a general LP (1).
- Through operations as described in Theorem 2 the program is changed into an equivalent standard program (2). They share the same optimal cost.
- The dual of program (D2) has the same optimal cost in turn.
- Both (D2) and (D1) have the same optimal cost by Theorem 2.
- Hence (1) and (D1) have the same optimal cost.

■

Complementary slackness

- Another important result that links both optima:

Theorem 5. *Let* $\mathbf{x}$ *and* μ *be feasible solutions to the primal and dual problems respectively. The vectors for* $\mathbf{x}$ *and* μ *are optimal solutions for the two respective problems if and only if*

$$\begin{aligned} u_i &= \mu_i(\mathbf{A}_i^T\mathbf{x} - b_i) && \mathbf{= 0}, \quad i = 1, .., m; \\ v_j &= (c_j - \mu^T\mathbf{a}_j)\mathbf{x}_j && \mathbf{= 0}, \quad j = 1, .., n. \end{aligned}$$

Proof. In the proof of the weak duality we showed that $u_i, v_j \geq 0$. Moreover

$$0 \leq \sum_i^m u_i + \sum_j^n v_j = \mathbf{c}^T\mathbf{x} - \mu^T\mathbf{b}.$$

Hence, $\mathbf{x}, \mu$ optimal $\Leftrightarrow u_i = v_j = 0$ through strong duality ($\Rightarrow$) and the second corollary of weak duality ($\Leftarrow$). ∎

Examples for LP's

Duality

- A simple example with the following linear program:

$$\begin{array}{ll} \text{minimize} & 3x_1 + x_2 \\ \text{subject to} & x_2 - 2x_1 = 1 \\ & x_1, x_2 \geq 0 \end{array}$$

- Two inequality constraints, one equality constraint. The Lagrangian is written:

$$L(x, \lambda, \mu) = 3x_1 + x_2 - \lambda_1 x_1 - \lambda_2 x_2 + \mu(1 - x_2 + 2x_1)$$

in the (dual variables) $\lambda_1, \lambda_2 \geq 0$ and μ (free).

Duality

- The dual function is then:

$$\begin{aligned} g(\lambda, \mu) &= \inf_{\mathbf{x}} L(\mathbf{x}, \lambda, \mu) \\ &= \inf_{\mathbf{x}} 3x_1 + x_2 - \lambda_1 x_1 - \lambda_2 x_2 + \mu(1 - x_2 + 2x_1) \\ &= \inf_{\mathbf{x}} (3 - \lambda_1 + 2\mu) x_1 + (1 - \lambda_2 - \mu) x_2 + \mu \end{aligned}$$

- We minimize a linear function of $x_1, \ x_2$, only two possibilities:

$$g(\lambda, \mu) = \begin{cases} \mu & \text{if } 3 - \lambda_1 + 2\mu = 1 - \lambda_2 - \mu = 0 \\ -\infty & \text{otherwise} \end{cases}$$

- The dual problem is finally:

$$\begin{array}{ll} \text{maximize} & \mu \\ \text{subject to} & 3 - \lambda_1 + 2\mu = 0 \\ & 1 - \lambda_2 - \mu = 0 \\ & \lambda \geq 0 \end{array}$$

Network flow: Max-flow / Min-cut

- m nodes, $N_1, \cdots, N_m$.
- d directed edges (arrows) to connect some nodes. Each edge is a pair $(N_i, N_{i'})$. The set is $\mathcal{V}$
  - Each edge carries a flow f_j its flow.
  - Each edge has a bounded capacity (pipe width) $f_j \leq u_j$
- Relating edges and nodes: the network's incidence matrix $A \in \{-1, 0, 1\}^{m \times d}$:

$$A_{ij} = \begin{cases} 1 & \text{if edge } j \text{ starts at node } i \\ -1 & \text{if edge } j \text{ ends at node } i \\ 0 & \text{otherwise} \end{cases}$$

- For a node i,

$$\sum_{j \text{ s.t. edge ends at } i} f_j = \sum_{j \text{ s.t. edge starts at } i} f_j$$

- In matrix form: $A\mathbf{f} = \mathbf{0}$

First problem: Maximal Flow

- We consider a **constant flow** from node 1 to node m.
- What is the **maximal flow** that can go through the system?
- A way to model this is to **close the loop** with an *artificial edge* numbered $d+1$.
- if $u_{d+1} = \infty$, what would be the maximal flow f_{d+1} of that edge?
- Namely solve

$$
\begin{array}{ll}
\text{minimize} & \mathbf{c}^T\mathbf{f} = -f_{d+1}, \\
\text{subject to} & [A\ , e]\,\mathbf{f} = 0, \\
& 0 \leq f_1 \leq u_1, \cdots, 0 \leq f_d \leq u_d, \\
& 0 \leq f_{d+1} \leq u_{d+1},
\end{array}
$$

with $e = (-1, 0, \ldots, 0, 1)$ and $c = (0, \ldots, 0, -1)$ and u_{d+1} a very large capacity for f_{d+1}.

Second problem: Minimal Cut

- Suppose you are a plumber and you want to completely **stop the flow** from node N_1 to N_m.
- You have to remove edges (pipes). What the minimal capacity you have to remove to make sure no flow goes from N_1 to N_m?
- Goal: cut the set of nodes into two disjoint sets S and T.
- Suppose we remove a set $\mathcal{C} \subset \mathcal{V}$ of edges. We want to minimize the total capacity of $\mathcal{C}$ under the constraint that the flow is now zero.
- $y_{ij} \in \{0, 1\}$ will keep track of cuts. For each node N_i there is a variable z_i which is 0 if N_i is in the set S or 1 in the set T. We arbitrarily set $z_1 = 0$ and $z_N = 1$.

$$\begin{array}{ll} \text{minimize} & \sum_{(i,j)\in\mathcal{V}} y_{ij}u_{ij} \\ \text{subject to} & y_{i,j} + z_i - z_j \geq 0 \\ & z_1 = 1, z_t = 0, z_i \geq 0, \\ & y_{ij} \geq 0, (i,j) \in \mathcal{V} \end{array}$$

Duality: example

- Let us form the **Lagrangian**:

$$L(\mathbf{f},\mathbf{y},\mathbf{z}) = \mathbf{c}^T\mathbf{f} + \mathbf{z}^T\,[Ae]\,\mathbf{f} + \mathbf{y}^T(\mathbf{f}-\mathbf{u})$$

  for $\mathbf{f} \geq 0$ here.

- The **Lagrange dual function** is defined as

$$\begin{aligned} g(\mathbf{y},\mathbf{z}) &= \inf_{\mathbf{f}\geq 0} L(\mathbf{f},\mathbf{y},\mathbf{z}) \\ &= \inf_{\mathbf{f}\geq 0} \mathbf{f}^T\left(\mathbf{c}+\mathbf{y}+\begin{bmatrix} A^T \\ e^T \end{bmatrix}\mathbf{z}\right) - \mathbf{u}^T\mathbf{y} \end{aligned}$$

- but this minimization yields either $-\infty$ or $-\mathbf{u}^T\mathbf{y}$, so:

$$g(\mathbf{y},\mathbf{z}) = \begin{cases} -\mathbf{u}^T\mathbf{y} & \text{if } \left(\mathbf{c}+\mathbf{y}+\begin{bmatrix} A^T \\ e^T \end{bmatrix}\mathbf{z}\right) \geq 0 \\ -\infty & \text{otherwise} \end{cases}$$

Duality: example

This means that the **dual** of the maximum flow problem is written:

$$\begin{array}{ll} \text{minimize} & \mathbf{u}^T \mathbf{y} \\ \text{subject to} & \mathbf{c} + \mathbf{y} + \begin{bmatrix} A^T \\ e \end{bmatrix} \mathbf{z} \geq 0 \end{array}$$

Compare the following dual with changed notations

$$\begin{array}{ll} \text{minimize} & \displaystyle\sum_{(i,j)\in\mathcal{V}} y_{ij} u_{ij} \\ \text{subject to} & y_{N,1} + z_N - z_1 \geq 1 \\ & y_{ij} + z_i - z_j \geq 0, \quad (i,j) \in \mathcal{V} \\ & y_{ij} \geq 0 \end{array}$$

to the **minimum cut problem**. The two problems are **identical**.

Duality: example

- The objective is to minimize:

$$\sum_{(i,j)\in\mathcal{V}} u_{ij}y_{ij}, \quad (y_{i,j} \geq 0),$$

where $u_{d+1} = u_{N,1} = M$ (very large), which means $y_{N,1} = 0$.

- The first equation then becomes:

$$z_N - z_1 \geq 1$$

so we can fix $z_N = 1$ and $z_1 = 0$.

Duality: example

- The equations for all the edges starting from $z_1 = 0$:

$$y_{1j} - z_j \geq 0$$

- Then, two scenarios are possible (no proof here):
  - $y_{1j} = 1$ with $z_j = 1$ and all the following z_k will be ones in the next equations (at the minimum cost):

$$y_{jk} + z_j - z_k \geq 0, \quad (j,k) \in \mathcal{V}$$

  - $y_{1j} = 0$ with $z_j = 0$ and we get the same equation for the next node:

$$y_{jk} - z_k \geq 0, \quad (j,k) \in \mathcal{V}$$

Duality: example

Interpretation?

- If a node has $z_i = 0$, all the nodes preceding it in the network must have $z_j = 0$.
- If a node has $z_i = 1$, all the following nodes in the network must have $z_j = 1 \ldots$
- This means that z_j effectively splits the network in two partitions
- The equations:
$$y_{ij} - z_i + z_j \geq 0$$
mean for any two nodes with $z_i = 0$ and $z_j = 1$, we must have $y_{ij} = 1$.
- The objective minimizes the total capacity of these edges, which is also the capacity of the cut.

Next time

- Geometric viewpoint on duality
- Sensitivity Analysis.