Today

Quickly: Matrix View, Taking Dual.

Matching, algebra, geometry.

Facility Location.

Rules for School...

or..."Rules for taking duals"
Canonical Form.

Primal LP	Dual LP
$\max c \cdot x$	$\min y^T b$
$Ax \leq b$	$y^T A \geq c$
$x \geq 0$	$y \geq 0$

Standard:

$Ax \leq b, \max cx, x \geq 0 \leftrightarrow y^T A \geq c, \min by, y \geq 0$.

min $\leftrightarrow$ max

$\geq \leftrightarrow \leq$

"inequalities" $\leftrightarrow$ "nonnegative variables"

"nonnegative variables" $\leftrightarrow$ "inequalities"

Another useful trick: Equality constraints. "equalities" $\leftrightarrow$ "unrestricted variables."

Maximum Weight Matching.

Bipartite Graph $G = (V, E)$, $w : E \rightarrow Z$.
Find maximum weight perfect matching.
Solution: x_e indicates whether edge e is in matching.

$$\max \sum_e w_e x_e$$
$$\forall v: \sum_{e=(u,v)} x_e = 1 \qquad p_v$$
$$x_e \geq 0$$

Dual.
Variable for each constraint. p_v unrestricted.
Constraint for each variable. Edge e, $p_u + p_v \geq w_e$
Objective function from right hand side. $\min \sum_v p_v$

$$\min \sum_v p_v$$
$$\forall e = (u,v): \quad p_u + p_v \geq w_e$$

Weak duality? Price function upper bounds matching.
$\sum_{e \in M} w_e x_e \leq \sum_{e=(u,v) \in M} p_u + p_v \leq \sum_v p_u$.
Strong Duality? Same value solutions. Hungarian algorithm !!!

Matrix View.

x_e variable for $e = (u, v)$.

			x_e		rhs
	·	⋯	0	⋯	1
	⋮	⋮	⋮	⋮	1
p_u	·	⋯	1	⋯	1
	·	⋯	0	⋯	1
	⋮	⋮	⋮	⋮	1
	·	⋯	0	⋯	1
p_v	·	⋯	1	⋯	1
	·	⋯	0	⋯	1
	⋮	⋮	⋮	⋮	1
obj	·	·	w_e	·	

Row equation: $\sum_{e=(u,v)} x_e = 1$. Row (dual) variable: p_u.

Column variable: x_e. Column (dual) constraint: $p_u + p_v \geq 1$.

Exercise: objectives?

Complementary Slackness.

$$\max \sum_e w_e x_e$$
$$\forall v: \sum_{e=(u,v)} x_e = 1 \qquad p_v$$
$$x_e \geq 0$$

Dual:
$$\min \sum_v p_v$$
$$\forall e = (u,v): \quad p_u + p_v \geq w_e$$

Complementary slackness:
Only match on tight edges.
Nonzero p_u on matched u.

Multicommodity Flow.

Given $G = (V, E)$, and capacity function $c : E \rightarrow Z$, and pairs $(s_1, t_1), \ldots, (s_k, t_k)$ with demands $d_1, \ldots, d_k$.
Route D_i flow for each s_i, t_i pair, so every edge has $\leq \mu c(e)$ flow with minimum μ.

variables: f_p flow on path p.
P_i -set of paths with endpoints s_i, t_i.

$$\min \mu$$
$$\forall e: \sum_{p \ni e} f_p \leq \mu c_e$$
$$\forall i: \sum_{p \in P_i} f_p = D_i$$
$$f_p \geq 0$$

Take the dual.

$$\min \mu$$
$$\forall e: \sum_{p \ni e} f_p \leq \mu c_e$$
$$\forall i: \sum_{p \in P_i} f_p = D_i$$
$$f_p \geq 0$$

Modify to make it $\geq$, which "go with min.
And only constants on right hand side.

$$\min \mu$$
$$\forall e: \mu c_e - \sum_{p \ni e} f_p \geq 0$$
$$\forall i: \sum_{p \in P_i} f_p = D_i$$
$$f_p \geq 0$$

Dual.

$$\min \mu$$
$$\forall e: \mu c_e - \sum_{p \ni e} f_p \geq 0 \qquad d_e$$
$$\forall i: \sum_{p \in P_i} f_p = D_i \qquad d_i$$
$$f_p \geq 0$$

Introduce variable for each constraint.
Introduce constraint for each var:
$\mu \rightarrow \sum_e c_e d_e = 1$. $f_p \rightarrow \forall p \in P_i\ d_i - \sum_{e \in p} d_e \leq 0$.
Objective: right hand sides. $\max \sum_i D_i d_i$

$$\max \sum_i D_i d_i$$
$$\forall p \in P_i: d_i \leq \sum_{e \in p} d(e) \qquad \sum_e c_e d_e = 1$$

d_i - shortest s_i, t_i path length. Toll problem!
Weak duality: toll lower bounds routing.
Strong Duality. Tight lower bound. First lecture. Or Experts.
Complementary Slackness: only route on shortest paths
only have toll on congested edges.

Matrix View

f_p variable for path $e_1, e_2, \ldots, e_k$. p connects s_i, t_i.

			f_p		μ	rhs
	$\cdot$	$\cdots$	0	$\cdots$	$\cdot$	0
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	0
d_{e_1}	$\cdot$	$\cdots$	-1	$\cdots$	c_{e_1}	0
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	0
d_{e_2}	$\cdot$	$\cdots$	-1	$\cdots$	c_{e_2}	0
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	0
d_{e_k}	$\cdot$	$\cdots$	-1	$\cdots$	c_{e_k}	0
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	0
d_i	$\cdot$	$\cdot$	1	$\cdot$	$\cdots$	D_i
obj	1	1	1	1		

Row constraint: $c_e \mu - \sum_{p \ni e} f_p \geq 0$. Row (dual) variable: d_e.

Row constraint: $\sum_{p \in P_i} f_p = D_i$. Row (dual) variable: d_i.

Column variable: f_p. Column (dual) constraint: $d_i - \sum_{e \in p} d_e \leq 0$.

Column variable: μ. Column (dual) constraint: $\sum_e d(e) c(e) = 1$.

Exercise: objectives?

Exponential size.

Multicommodity flow.

$$\min \mu$$
$$\forall e: \mu c_e - \sum_{p \ni e} f_p \geq 0$$
$$\forall i: \sum_{p \in P_i} f_p = d_i$$
$$f_p \geq 0$$

Dual is.

$$\max \sum_i D_i d_i$$
$$\forall p \in P_i: d_i \leq \sum_{e \in p} d(e)$$

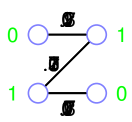
Exponential sized programs?

Answer 1: We solved anyway!

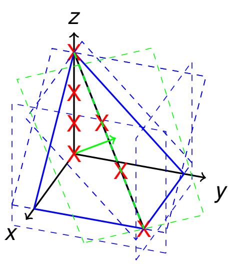
Answer 2: Ellipsoid algorithm.
Find violated constraint $\rightarrow$ poly time algorithm.

Answer 3: there is polynomial sized formulation.
Question: what is it?

Maximum matching and simplex.

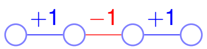


$$\max x + y + z$$
$$x \leq 1$$
$$x + z \leq 1 \qquad a = 1$$
$$z + y \leq 1 \qquad b = 1$$
$$y \leq 1$$
$$x \geq 0$$
$$y \geq 0$$
$$z \geq 0 \qquad c = 1$$



Blue constraints tight.

Sum: $x + 2z + y$.



Augmenting Path. Via Gaussian Elimination!

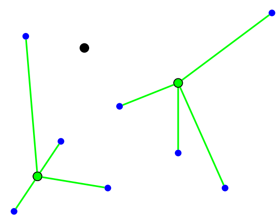
Facility location

Set of facilities: F, opening cost f_i for facility i

Set of clients: D.

d_{ij} - distance between i and j.
(notation abuse: clients/facility confusion.)

Triangle inequality: $d_{ij} \leq d_{ik} + d_{kj}$.



Facility Location

Linear program relaxation:

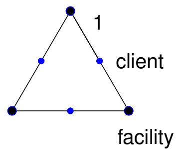
"Decision Variables".
y_i - facility i open?
x_{ij} - client j assigned to facility i.

$$\min \sum_{i\in F} y_i f_i + \sum_{i\in F, j\in D} x_{ij} d_{ij}$$
$$\forall j \in D \quad \sum_{i\in F} x_{ij} \geq 1$$
$$\forall i \in F, j\in D \quad x_{ij} \leq y_i,$$
$$x_{ij}, y_i \geq 0$$

Facility opening cost.
Client Connnection cost.
Must connect each client.
Only connect to open facility.

Integer Solution?

$$\min \sum_{i\in F} y_i f_i + \sum_{i\in F, j\in D} x_{ij} d_{ij}$$
$$\forall j \in D \quad \sum_{i\in F} x_{ij} \geq 1$$
$$\forall i \in F, j\in D \quad x_{ij} \leq y_i,$$
$$x_{ij}, y_i \geq 0$$



$x_{ij} = \frac{1}{2}$ edges.
$y_i = \frac{1}{2}$ edges.

Facility Cost: $\frac{3}{2}$ Connection Cost: 3
Any one Facility:
Facility Cost: 1 Client Cost: 3.7
Make it worse? Sure. Not as pretty!

Round solution?

$$\min \sum_{i\in F} y_i f_i + \sum_{i\in F, j\in D} x_{ij} d_{ij}$$
$$\forall j \in D \quad \sum_{i\in F} x_{ij} \geq 1$$
$$\forall i \in F, j\in D \quad x_{ij} \leq y_i,$$
$$x_{ij}, y_i \geq 0$$

Round independently?

y_i and x_{ij} separately? Assign to closed facility!

Round x_{ij} and open facilities?
Different clients force different facilities open.

Any ideas?

Use Dual!

The dual.

$\min cx, Ax \geq b \leftrightarrow \max bx, y^T A \leq c$.

$$\min \sum_{i\in F} y_i f_i + \sum_{i\in F, j\in D} x_{ij} d_{ij}$$
$$\forall j \in D \quad \sum_{i\in F} x_{ij} \geq 1$$
$$\forall i \in F, j\in D \quad x_{ij} \leq y_i,$$

$$\min \sum_{i\in F} y_i f_i + \sum_{i\in F, j\in D} x_{ij} d_{ij}$$
$$\forall j \in D \quad \sum_{i\in F} x_{ij} \geq 1 \quad ; \alpha_j$$
$$\forall i \in F, j\in D \quad y_i - x_{ij} \geq 0 \quad ; \beta_{ij}$$
$$x_{ij}, y_i \geq 0$$

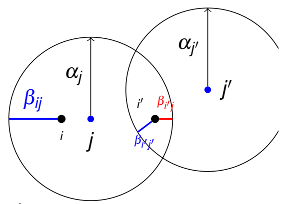
$$\max \sum_j \alpha_j$$
$$\forall i \quad \sum_{j\in D} \beta_{ij} \leq f_i \quad ; y_i$$
$$\forall i \in F, j\in D \quad \alpha_i - \beta_{ij} \leq d_{ij} \quad ; x_{ij}$$
$$\beta_{ij}, \alpha_j \geq 0$$

Interpretation of Dual?

$$\min \sum_{i\in F} y_i f_i + \sum_{i\in F, j\in D} x_{ij} d_{ij}$$
$$\forall j \in D \quad \sum_{i\in F} x_{ij} \geq 1$$
$$\forall i \in F, j\in D \quad x_{ij} \leq y_i,$$
$$x_{ij}, y_i \geq 0$$

$$\max \sum_j \alpha_j$$
$$\forall i \in F \quad \sum_{j\in D} \beta_{ij} \leq f_i$$
$$\forall i \in F, j\in D \quad \alpha_j - \beta_{ij} \leq d_{ij} \quad x_{ij}$$
$$\alpha_j, \beta_{ij} \leq 0$$

α_j charge to client.
maximize price paid by client to connect!
Objective: $\sum_j \alpha_j$ total payment.
Client j travels or pays to open facility i.
Costs client d_{ij} to get to there.
Savings is $\alpha_j - d_{ij}$.
Willing to pay $\beta_{ij} = \alpha_j - d_{ij}$.
Total payment to facility i at most f_i before opening.
Complementary slackness: $x_{ij} \geq 0$ if and only if $\alpha_j \geq d_{ij}$.
only assign client to "paid to" facilities.



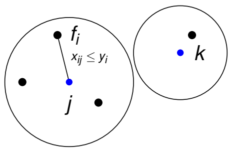
Use Dual.

1. Find solution to primal, (x,y). and dual, (α,β).
2. For smallest (remaining) α_j,
   (a) Let $N_j = \{i : x_{ij} > 0\}$.
   (b) Open cheapest facility i in N_j.
   Every client j' with $N_{j'} \cap N_j \neq \emptyset$ assigned to i.
3. Removed assigned clients, goto 2.

Integral facility cost at most LP facility cost.

Claim: Total facility cost is at most $\sum_i f_i y_i$.

2. For smallest (remaining) α_j,
   (a) Let $N_j = \{i : x_{ij} > 0\}$.
   (b) Open cheapest facility i in N_j.
   Every client j' with $N_{j'} \cap N_j \neq \emptyset$ assigned to i.



Proof: Step 2 picks client j.
$f_{\min}$ - min cost facility in N_j

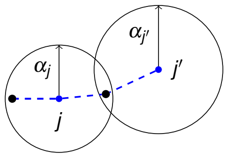
$f_{\min} \leq f_{\min} \cdot \sum_{i \in N_j} x_{ij} \leq f_{\min} \sum_{i \in N_j} y_i \leq \sum_{i \in N_j} y_i f_i$.

For k used in Step 2.
$N_j \cap N_k = \emptyset$ for j and k in step 2.
$\rightarrow$ Any facility in ≤ 1 sum from step 2.
$\rightarrow$ total step 2 facility cost is $\sum_i y_i f_i$.

Connection Cost.

2. For smallest (remaining) α_j,
   (a) Let $N_j = \{i : x_{ij} > 0\}$.
   (b) Open cheapest facility i in N_j.
   Every client j' with $N_{j'} \cap N_j \neq \emptyset$ assigned to i.

Client j is directly connected. Clients j' are indirectly connected.



Connection Cost of j: $\leq \alpha_j$.
Connection Cost of j':
$\leq \alpha_{j'} + \alpha_j + \alpha_j \leq 3\alpha_{j'}$.
since $\alpha_j \leq \alpha_{j'}$

Total connection cost:
at most $3\sum_{j'} \alpha_j \leq 3$ times Dual OPT.

Previous Slide: Facility cost:
$\leq$ primal "facility" cost $\leq$ Primal OPT.

Total Cost: 4 OPT.

Twist on randomized rounding.

Client j: $\sum_i x_{ij} = 1$, $x_{ij} \geq 0$.
Probability distribution! $\rightarrow$ Choose from distribution, x_{ij}, in step 2.

Expected opening cost:
$\sum_{i \in N_j} x_{ij} f_i \leq \sum_{i \in N_j} y_i f_i$.
and separate balls implies total $\leq \sum_i y_i f_i$.

$D_j = \sum_i x_{ij} d_{ij}$ Connection cost of primal for j.

Expected connection cost j' $\alpha_j + \alpha_{j'} + D_j$.

In step 2: pick in increasing order of $\alpha_j + D_j$.

$\rightarrow$ Expected cost is $(2\alpha_{j'} + D_{j'})$. Connection cost: $2\sum_j \alpha_j + \sum_j D_j$.
$2OPT(D)$ plus connection cost or primal.

Total expected cost:
Facility cost is at most facility cost of primal.
Connection cost at most $2OPT$ + connection cost of prmal.
$\rightarrow$ at most $3OPT$.

Primal dual algorithm.

1. Feasible integer solution.
2. Feasible dual solution.
3. Cost of integer solution $\leq \alpha$ times dual value.

Just did it. Used linear program. Faster?

Typically.
Begin with feasible dual.
Raise dual variables until tight constraint.
Set corresponding primal variable to an integer.

Recall Dual:

$$\max \sum_j \alpha_j$$
$$\forall i \in F \quad \sum_{j \in D} \beta_{ij} \leq f_i$$
$$\forall i \in F, j \in D \quad \alpha_j - \beta_{ij} \leq d_{ij}$$
$$\alpha_j, \beta_{ij} \leq 0$$

Facility location primal dual.

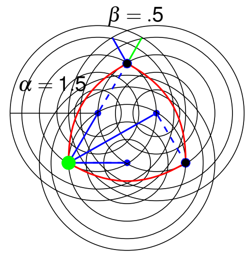
Phase 1: 1. Initially $\alpha_j, \beta_{ij} = 0$.
2. Raise α_j for every (unconnected) client.
When $\alpha_j = d_{ij}$ for some i
raise β_{ij} at same rate Why? Dual: $\alpha_j - \beta_{ij} \leq d_{ij}$.
Intution:Paying β_{ij} to open i.
Stop when $\sum_j \beta_{ij} = f_i$.
Why? Dual: $\sum_i \beta_{ij} \leq f_i$
Intution: facility paid for.
Temporarily open i.
Connect all tight ji clients j to i.

3. Continue until all clients connected.

Phase 2:
Connect facilities that were paid by same client.
Permanently open an independent set of facilities.

For client j, connected facility i is opened. Good.
Connected facility not open
$\rightarrow$ exists client j' paid i and connected to open facility.

Connect j to j''s open facility.



Constraints for dual.
$\sum_j \beta_{ij} \leq f_i$
$\alpha_i - \beta_{ij} \leq d_{ij}$.
Grow α_j.
$\alpha_j = d_{ij}$!
Tight constraint: $\alpha_j - \beta_{ij} \leq d_{ij}$.
Grow β_{ij} (and α_j).
$\sum_j \beta_{ij} = f_i$ for all facilities.
Tight: $\sum_j \beta_{ij} \leq f_i$
LP Cost: $\sum_j \alpha_j = 4.5$

Temporarily open all facilities.
Assign Clients to "paid to" open facility.
Connect facilities with client that pays both.
Open independent set.
Connect to "killer" client's facility.
Cost: 1 + 3.7 = 4.7.
A bit more than the LP cost.

Analysis

Claim: Client only pays one facility.

Independent set of facilities.

Claim: S_i - directly connected clients to open facility i.
$f_i + \sum_{j \in S_i} d_{ij} \leq \sum_j \alpha_j$.

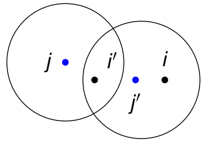
Proof:
$f_i = \sum_{j \in S_i} \beta_{ij} = \sum_{j \in S_i} \alpha_j - d_{ij}$.
Since directly connected: $\beta_{ij} = \alpha_j - d_{ij}$. □

See you on Thursday.

Analysis.

Claim: Client j is indirectly connected to i
$\rightarrow d_{ij} \leq 3\alpha_j$.

Directly connected to (temp open) i'
conflicts with i.
exists j' with $\alpha_{j'} \geq d_{ij'}$ and $\alpha_j \geq d_{i'j'}$.
When i' opens, stops both α_j and α'_j.
α'_j stopped no later (..maybe earlier..)
$\alpha_j \leq \alpha'_j$.
Total distance from j to i'.
$d_{ij} + d_{ij'} + d_{i'j'} \leq 3\alpha_j$ □



Putting it together!

Claim: Client only pays one facility.
Claim: S_i - directly connected clients to open facility i.
$f_i + \sum_{j \in S_i} d_{ij} \leq \sum_j \alpha_j$.
Claim: Client j is indirectly connected to i
$\rightarrow d_{ij} \leq 3\alpha_j$.

Total Cost:
direct clients dual (α_j) pays for facility and own connections.
plus no more than 3 times indirect client dual.
Total Cost: 3 times dual.
feasible dual upper bounds fractional (and integer) primal.

3 OPT.

Fast! Cheap! Safe!