Trasporto e assegnamento

Indice

Il problema di trasporto

Soluzioni di base

Soluzioni ammissibili di base iniziali

Il metodo dei potenziali

Il problema di assegnamento

Errori tipici

Obiettivi (1/2)

Alla fine di questo capitolo sarai in grado di:

- formulare il problema di trasporto e rappresentarlo in tabella, bilanciare un'istanza aggiungendo un'origine o una destinazione fittizia e ridurre un problema di trasbordo senza capacità a un problema di trasporto;
- interpretare le basi del modello di trasporto come alberi ricoprenti di una rete bipartita, calcolare una soluzione di base e riconoscerne la degenerazione;
- costruire una soluzione ammissibile di base iniziale con la regola dell'angolo nord-ovest o quella del costo minimo;

Obiettivi (2/2)

Alla fine di questo capitolo sarai in grado di:

- applicare il metodo dei potenziali, calcolare i costi ridotti e i cicli di pivot e certificare l'ottimalità con il modello duale;
- formulare il problema di assegnamento, dimostrare il teorema di König–Egerváry ed eseguire l'algoritmo ungherese, interpretando i potenziali come una soluzione duale ottima.

Prerequisiti

- Da questo modulo, Capitolo 1 «Grafi»: alberi, foglie, cicli fondamentali e rango della matrice di incidenza nodo-arco.
- Da questo modulo, Capitolo 3 «Cammini minimi»: distanze, e Capitolo 4 «Modelli di flusso su reti»: il modello di flusso di costo minimo, il nodo fittizio, vertici interi, potenziali e costi ridotti.
- Dal modulo Programmazione lineare, Capitolo 9 «Teoria della dualità»: il duale di un modello LP e i teoremi della dualità debole e forte.

Il problema di trasporto

Introduzione: Il problema di trasporto

- Merci disponibili in alcune origini devono essere spedite ad alcune destinazioni, a costo totale minimo
- Scriviamo un modello LP per questo problema e la sua tabella di trasporto, e vediamo che è un modello di flusso di costo minimo su una rete bipartita completa
- Vediamo quando un'istanza è ammissibile, come bilanciare un'istanza non bilanciata e come un'istanza di trasbordo senza capacità diventa un'istanza di trasporto

Definizione 1: problema di trasporto (1/2)

*Ci sono m **origini** e n **destinazioni**. Numeriamo le origini $1, 2, \ldots, m$ e le destinazioni $m+1, m+2, \ldots, m+n$, e scriviamo*

$$P = \{1, 2, \ldots, m\}, \qquad D = \{m+1, m+2, \ldots, m+n\}.$$

- *Ogni origine $i \in P$ ha un'**offerta** $a_i > 0$, ogni destinazione $j \in D$ ha una **domanda** $d_j > 0$, e $c_{ij} \in \mathbb{R}$ è il costo di spedizione di un'unità da i a j.*

Definizione: problema di trasporto (2/2)

- *L'istanza è* ***bilanciata*** *se* $\sum_{i \in P} a_i = \sum_{j \in D} d_j$.

Il ***problema di trasporto*** *chiede le quantità* $x_{ij} \geq 0$ *da spedire da ogni origine* i *a ogni destinazione* j*, in modo che ogni origine spedisca tutta la sua offerta e ogni destinazione riceva la sua domanda, a costo totale minimo.*

La rete e le variabili

- In questo capitolo m e n indicano il numero di origini e di destinazioni
- La rete associata ha $m + n$ nodi e un arco (i, j) per ogni $i \in P$ e $j \in D$, cioè $m\,n$ archi: è la **rete bipartita completa** di P e D, perché ogni arco va da P a D
- Numerare le destinazioni dopo le origini dà a ogni nodo un suo numero, e più avanti un suo potenziale
- Usiamo una variabile continua x_{ij} per ogni coppia $i \in P$, $j \in D$: la quantità spedita da i a j

Il modello TP

Con queste variabili, un modello LP per il problema di trasporto è il seguente:

$$\text{min} \quad \sum_{i \in P} \sum_{j \in D} c_{ij}\, x_{ij}$$

$$\text{soggetto a} \quad \sum_{j \in D} x_{ij} = a_i, \quad \forall i \in P,$$

$$\sum_{i \in P} x_{ij} = d_j, \quad \forall j \in D,$$

$$x_{ij} \geq 0, \quad \forall i \in P,\ \forall j \in D.$$

Questo modello LP è detto modello TP.

Il modello TP: funzione obiettivo e vincoli

Descrizione della funzione obiettivo e dei vincoli del modello TP:

- la funzione obiettivo minimizza il costo totale di spedizione;
- i vincoli di offerta, uno per ogni origine, spediscono tutta l'offerta dell'origine;
- i vincoli di domanda, uno per ogni destinazione, consegnano tutta la domanda della destinazione;
- infine, gli ultimi vincoli definiscono le variabili del modello LP.

Un modello di flusso di costo minimo

- Moltiplicando per -1 i vincoli di domanda si ottengono i vincoli di conservazione del flusso della rete bipartita completa, con bilanci

$$b_i = a_i \text{ per } i \in P, \qquad b_j = -d_j \text{ per } j \in D,$$

  e senza capacità

- Il modello TP è il modello MCF della riga «Trasporto» della tabella «Casi speciali» del Capitolo 4 «Modelli di flusso su reti»

Proposizione 1: ammissibilità e integralità

Un'istanza del problema di trasporto ha una soluzione ammissibile se e solo se è bilanciata. Un'istanza bilanciata ha una soluzione ottima e, se le offerte a_i e le domande d_j sono intere, ha una soluzione ottima intera.

Dimostrazione (1/2)

Sia $S = \sum_{i \in P} a_i$.

1. *Il bilanciamento è necessario.* Se $\boldsymbol{x}$ è ammissibile, sommando i vincoli di offerta e i vincoli di domanda si ottiene la stessa somma di tutte le variabili:

$$\sum_{i \in P} a_i = \sum_{i \in P} \sum_{j \in D} x_{ij} = \sum_{j \in D} d_j.$$

2. *Il bilanciamento è sufficiente.* Se $S = \sum_{j \in D} d_j$, le quantità $x_{ij} = a_i \, d_j / S \geq 0$ sono ammissibili: la riga i ha somma $a_i \sum_{j \in D} d_j / S = a_i$, e la colonna j ha somma $d_j \sum_{i \in P} a_i / S = d_j$.

Dimostrazione (2/2)

3. *Soluzione ottima.* Ogni soluzione ammissibile ha $0 \leq x_{ij} \leq a_i$, quindi la regione ammissibile è un poliedro limitato non vuoto, e una funzione lineare vi raggiunge il minimo in un vertice (Capitolo 15 «La geometria della Programmazione lineare» del modulo Programmazione lineare).

4. *Integralità.* Con bilanci interi e senza capacità, ogni vertice è intero per il Teorema «vertici interi» del Capitolo 4.

□

La tabella di trasporto

- La **tabella di trasporto** (tableau) è una tabella $m \times n$ con una riga per ogni origine e una colonna per ogni destinazione
- La cella (i, j) contiene il costo c_{ij}, o più avanti la quantità x_{ij}; l'ultima colonna contiene le offerte a_i e l'ultima riga le domande d_j
- Un vettore $\boldsymbol{x} \geq 0$ scritto nella tabella è ammissibile se e solo se ogni riga ha per somma la sua offerta e ogni colonna la sua domanda

Esempio: tre magazzini e quattro negozi (1/3)

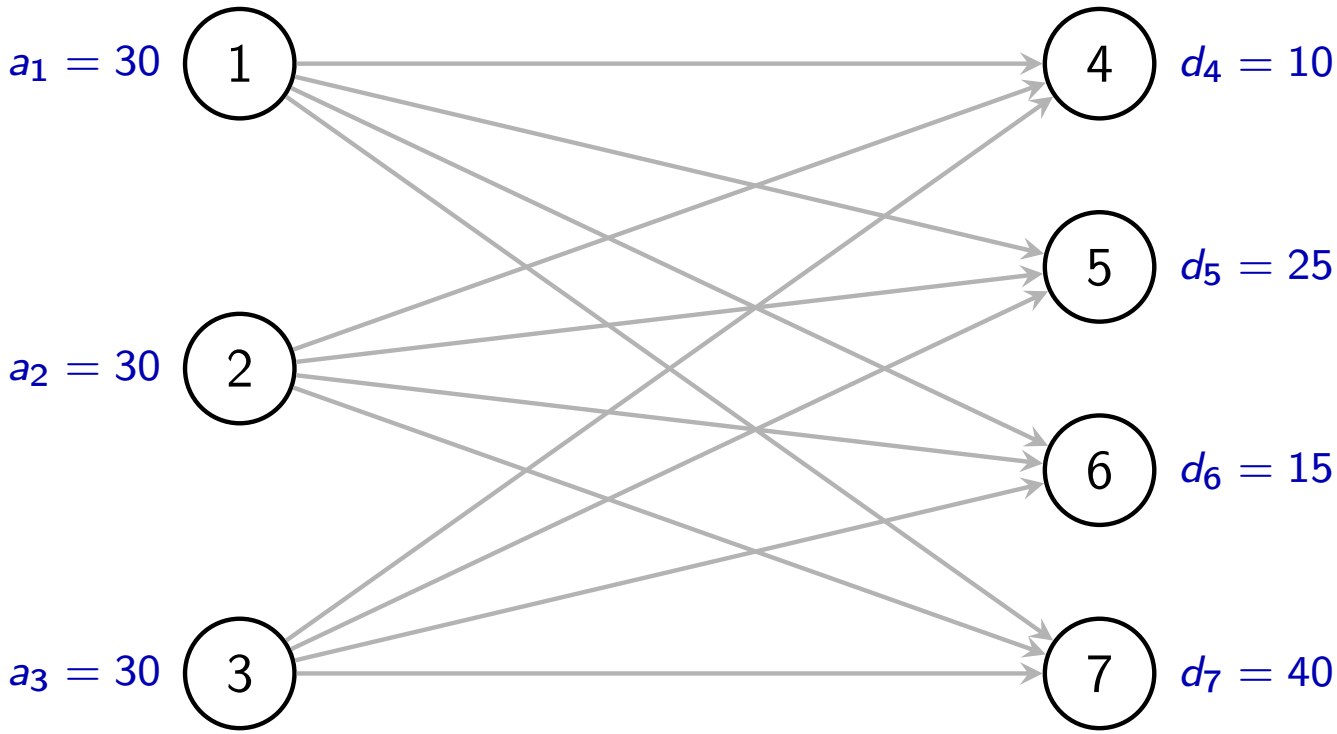
Un'azienda ha tre magazzini, le origini $1, 2, 3$, con 30 unità ciascuno, e quattro negozi, le destinazioni $4, 5, 6, 7$, con domande 10, 25, 15 e 40. I costi unitari sono nella tabella:

c_{ij}	4	5	6	7	a_i
1	1	9	2	3	30
2	3	4	2	2	30
3	9	1	4	6	30
d_j	10	25	15	40	90

L'istanza è bilanciata: l'offerta totale e la domanda totale valgono entrambe 90.

Esempio: tre magazzini e quattro negozi (2/3)

Nella rete bipartita completa ogni origine è collegata a ogni destinazione, con $3 \cdot 4 = 12$ archi; accanto a ogni nodo, la sua offerta o la sua domanda.



Esempio: tre magazzini e quattro negozi (3/3)

- Le quantità proporzionali $x_{ij} = a_i\, d_j/90$ della Proposizione «ammissibilità e integralità» sono ammissibili, per esempio

$$x_{14} = \frac{30 \cdot 10}{90} = \frac{10}{3},$$

  ma frazionarie

- I vertici della regione ammissibile sono interi

Usiamo questa istanza come esempio del problema di trasporto in tutto il capitolo.

Istanze non bilanciate: una destinazione fittizia

- Se l'offerta totale supera la domanda totale e ogni origine spedisce *al più* la sua offerta, aggiungiamo una **destinazione fittizia** con domanda

$$\sum_{i \in P} a_i - \sum_{j \in D} d_j$$

  e costo 0 da ogni origine

- La quantità spedita alla destinazione fittizia è l'offerta non usata dell'origine: è il nodo fittizio del Capitolo 4, come colonna della tabella

Istanze non bilanciate: un'origine fittizia

- Se la domanda totale supera l'offerta totale e una destinazione può ricevere meno della sua domanda, aggiungiamo un'**origine fittizia** con offerta

$$\sum_{j \in D} d_j - \sum_{i \in P} a_i$$

- Il costo della sua cella $(\cdot, j)$ è la penale per ogni unità di domanda di j non soddisfatta, oppure 0 se non c'è penale

Esempio: una destinazione fittizia

Nell'esempio dei tre magazzini l'offerta del magazzino 1 sale a $a_1 = 40$: l'offerta totale è 100 e la domanda totale 90.

- Aggiungiamo la destinazione fittizia 8 con domanda $d_8 = 100 - 90 = 10$, e le celle $(1, 8)$, $(2, 8)$, $(3, 8)$ con costo 0: la tabella ha una quinta colonna.

- La soluzione ottima, calcolata con Gurobi, è unica:

$$\tilde{x}_{14} = 10,\ \tilde{x}_{16} = 15,\ \tilde{x}_{17} = 10,\ \tilde{x}_{18} = 5,$$
$$\tilde{x}_{27} = 30,\ \tilde{x}_{35} = 25,\ \tilde{x}_{38} = 5,$$

  con costo 155.

- I magazzini 1 e 3 tengono 5 unità ciascuno, e il magazzino 2 spedisce tutta la sua offerta.

Il trasbordo come trasporto

- In un'istanza di trasbordo le unità possono passare per nodi intermedi
- Quando gli archi non hanno capacità, ogni unità può viaggiare lungo un cammino minimo, e contano soltanto le distanze fra i nodi di offerta e i nodi di domanda
- La proposizione che segue lo rende preciso: un'istanza di trasbordo senza capacità è un'istanza di trasporto con le distanze come costi

Proposizione 2: trasbordo come trasporto

Sia $G = (N, A)$ una rete con bilanci b_i, limiti inferiori nulli, senza capacità e con costi $c_{ij} \geq 0$. Siano $P = \{i \in N : b_i > 0\}$ e $D = \{j \in N : b_j < 0\}$, e supponiamo che ogni $j \in D$ sia raggiungibile da ogni $i \in P$, con distanza $\tilde{d}(i, j)$ per le lunghezze c_{ij}. Allora il costo minimo di un flusso ammissibile è uguale al valore ottimo dell'istanza di trasporto con origini P, offerte $a_i = b_i$, destinazioni D, domande $d_j = -b_j$ e costi $\tilde{d}(i, j)$.

Dimostrazione (1/4)

Confrontiamo i due valori ottimi.

1. *Dalle spedizioni ai flussi.* Sia $\boldsymbol{y}$ ammissibile per l'istanza di trasporto. Inviando y_{ij} unità lungo un cammino minimo da i a j, per ogni coppia, si ottiene un flusso $\boldsymbol{x} \geq 0$:
   - il nodo $i \in P$ invia $\sum_{j \in D} y_{ij} = b_i$ unità, il nodo $j \in D$ ne riceve $-b_j$, gli altri nodi inoltrano ciò che ricevono.

   Il suo costo è $\sum_{i,j} \tilde{d}(i,j)\, y_{ij}$, quindi il costo minimo di un flusso è al più l'ottimo del trasporto.

Dimostrazione (2/4)

2. *Dai flussi alle spedizioni.* Mostriamo per induzione sul numero di archi con $x_{ij} > 0$ più il numero di nodi con $b_i \neq 0$ che ogni flusso ammissibile $\boldsymbol{x}$ costa almeno $\sum_{i,j} \tilde{d}(i,j)\, y_{ij}$ per qualche $\boldsymbol{y}$ ammissibile. Se $\boldsymbol{b} = 0$, prendiamo $\boldsymbol{y} = 0$.

3. Altrimenti partiamo da un nodo i_0 con $b_{i_0} > 0$: il suo flusso uscente è positivo, quindi ha un arco uscente con flusso positivo. Un nodo v con $b_v \geq 0$ raggiunto lungo un tale arco ha flusso entrante positivo, quindi flusso uscente positivo. Seguiamo questi archi finché un nodo si ripete o si raggiunge un nodo j_0 con $b_{j_0} < 0$.

Dimostrazione (3/4)

4. Se un nodo si ripete, gli archi formano un ciclo orientato C con flussi positivi. Sottrarre il loro minimo ε lungo C mantiene il flusso ammissibile, non aumenta il costo, perché $c \geq 0$, e svuota un arco: si applica l'induzione.

5. Altrimenti abbiamo un cammino Q da i_0 a j_0. Sia $\varepsilon > 0$ il minimo fra b_{i_0}, $-b_{j_0}$ e i flussi su Q. Sottrarre ε lungo Q, e da b_{i_0} e $-b_{j_0}$, dà un flusso ammissibile $\boldsymbol{x}'$ per i nuovi bilanci, con meno archi positivi o bilanci non nulli, e

$$\boldsymbol{c}'\boldsymbol{x} = \boldsymbol{c}'\boldsymbol{x}' + \varepsilon\, c(Q) \geq \boldsymbol{c}'\boldsymbol{x}' + \varepsilon\, \tilde{d}(i_0, j_0).$$

Dimostrazione (4/4)

- Per induzione
$$\boldsymbol{c}'\boldsymbol{x}' \geq \sum \tilde{d}(i,j)\, y'_{ij}$$
- Aggiungendo ε a $y'_{i_0 j_0}$ si ottiene il $\boldsymbol{y}$ cercato:
$$\boldsymbol{c}'\boldsymbol{x} \geq \sum \tilde{d}(i,j)\, y'_{ij} + \varepsilon\, \tilde{d}(i_0,j_0) = \sum \tilde{d}(i,j)\, y_{ij}$$

□

Esempio: trasbordo senza capacità (1/3)

Prendiamo la rete di riferimento del Capitolo 4, con gli stessi bilanci $\boldsymbol{b} = (5, 2, 0, -3, 0, -4)$ e gli stessi costi, e togliamo le capacità.

- Le origini sono i nodi di offerta $1, 2$ e le destinazioni i nodi di domanda $4, 6$; teniamo i nomi della rete.
- Le distanze sono

$$\tilde{d}(1,4) = 7 \text{ lungo } 1,2,4, \quad \tilde{d}(1,6) = 8 \text{ lungo } 1,3,5,6,$$

$$\tilde{d}(2,4) = 5 \text{ lungo } 2,4, \quad \tilde{d}(2,6) = 10 \text{ lungo } 2,4,6.$$

Esempio: trasbordo senza capacità (2/3)

- L'istanza di trasporto 2×2 ha $a_1 = 5$, $a_2 = 2$, $d_4 = 3$, $d_6 = 4$
- Con $y_{24} = t \in [0, 2]$ le altre quantità sono $y_{14} = 3 - t$, $y_{16} = 2 + t$, $y_{26} = 2 - t$, e il costo è

$$7(3 - t) + 8(2 + t) + 5t + 10(2 - t) = 57 - 4t$$

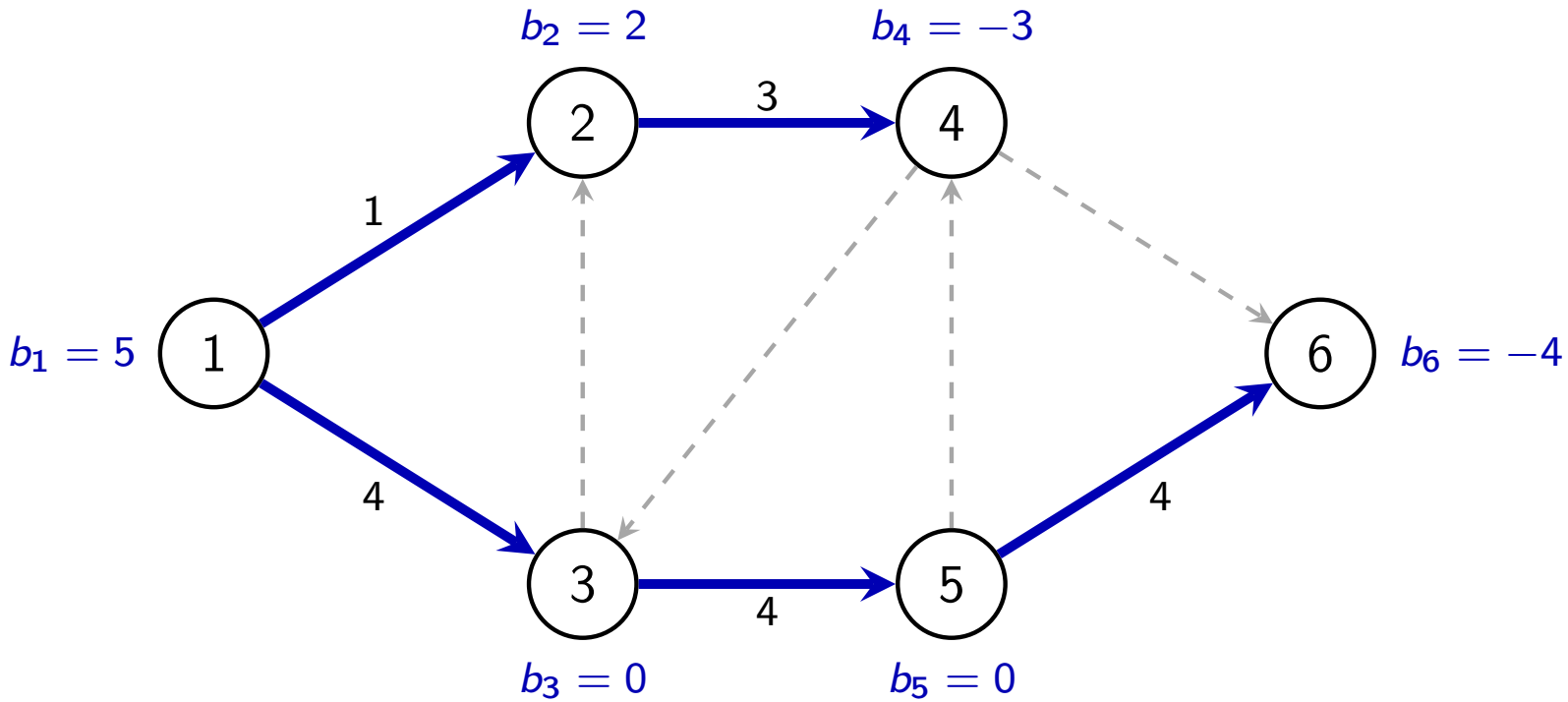
- Il minimo si ha per $t = 2$:

$$y_{14} = 1, \quad y_{16} = 4, \quad y_{24} = 2, \quad y_{26} = 0,$$

con costo 49

Esempio: trasbordo senza capacità (3/3)

Lungo i cammini minimi si ottiene il flusso $x_{12} = 1$, $x_{13} = 4$, $x_{24} = 3$, $x_{35} = 4$, $x_{56} = 4$ (blu), di costo 49, l'ottimo calcolato da Gurobi.



Con le capacità del Capitolo 4 questo flusso non è ammissibile, perché l'arco $(5, 6)$ ha capacità 2, e l'ottimo è 60.

Riepilogo

Il problema di trasporto:

- spedisce le offerte di m origini alle domande di n destinazioni a costo minimo
- è un modello di flusso di costo minimo sulla rete bipartita completa

Ammissibilità:

- un'istanza è ammissibile esattamente quando è bilanciata
- un'origine o una destinazione fittizia la bilancia

Trasbordo:

- senza capacità, è un'istanza di trasporto con le distanze come costi

Soluzioni di base

Introduzione: Soluzioni di base

- Le soluzioni ottime di un modello LP si trovano fra i suoi vertici, cioè fra le sue soluzioni ammissibili di base
- In questa sezione vediamo che le basi del modello TP hanno una forma semplice: sono gli alberi ricoprenti della rete bipartita completa
- È la chiave del metodo dei potenziali: una soluzione di base si calcola foglia per foglia, direttamente nella tabella

La matrice dei vincoli

- La colonna di x_{ij} nei vincoli del modello TP ha due elementi uguali a 1, nella riga dell'offerta di i e nella riga della domanda di j
- Cambiando il segno delle righe di domanda si ottiene la matrice di incidenza $\boldsymbol{B}$ della rete bipartita completa
- Questo non cambia né il rango né le dipendenze lineari delle colonne

Proposizione 3: rango del modello di trasporto

Gli $m + n$ vincoli del modello TP hanno rango $m + n - 1$. In un'istanza bilanciata ogni vincolo è implicato dagli altri.

Dimostrazione

- Ogni origine è collegata a ogni destinazione, quindi la rete bipartita completa è connessa, e la sua matrice di incidenza ha rango $m + n - 1$ per il Teorema «rango della matrice di incidenza» del Capitolo 1.

- La somma delle righe di offerta meno la somma delle righe di domanda è nulla, perché entrambe le somme contengono ogni variabile una volta, e la stessa combinazione dei termini noti è

$$\sum_{i \in P} a_i - \sum_{j \in D} d_j = 0 :$$

ogni vincolo è combinazione degli altri.

□

Definizione 2: base e soluzione di base

*Una **base** del modello* TP *è un insieme T di $m+n-1$ celle i cui archi formano un albero ricoprente della rete bipartita completa; le sue celle sono le **celle di base**.*

*Una **soluzione di base** di T è una soluzione dei vincoli di offerta e di domanda con $x_{ij} = 0$ per ogni cella fuori da T.*

*È una **soluzione ammissibile di base** se $\mathbf{x} \geq 0$, ed è **degenere** se qualche cella di base ha $x_{ij} = 0$.*

Teorema 1: le basi sono alberi ricoprenti

(a) *Un insieme di $m + n - 1$ celle ha colonne linearmente indipendenti se e solo se i suoi archi formano un albero ricoprente.*

(b) *In un'istanza bilanciata, ogni base T ha esattamente una soluzione di base. Si trova fissando, una dopo l'altra, la cella di una foglia dell'albero, ed è intera quando le offerte e le domande sono intere.*

Dimostrazione (1/3)

- La parte (a) è il Corollario «basi della matrice di incidenza» del Capitolo 1, perché il segno delle righe di domanda non cambia le dipendenze delle colonne.

Per la parte (b):

1. *Unicità.* Due soluzioni di base di T differiscono per un vettore $\boldsymbol{y}$ nullo fuori da T che soddisfa i vincoli con termini noti nulli; le colonne di T sono indipendenti per (a), quindi $\boldsymbol{y} = 0$.

Dimostrazione (2/3)

2. *Costruzione.* L'albero ha una foglia v (Lemma «foglie» del Capitolo 1), con una sola cella e di T. Il vincolo di v non contiene altre celle di base, quindi x_e è uguale all'offerta o alla domanda di v.

3. Togliamo v ed e, e sottraiamo x_e dall'offerta o dalla domanda dell'altro estremo di e. Le celle rimaste formano un albero ricoprente dei nodi rimasti, e le offerte e domande rimaste hanno ancora somme uguali.

Dimostrazione (3/3)

4. Quando resta una sola cella, entrambi i suoi estremi sono foglie, e le somme uguali danno lo stesso valore per i due vincoli.

   Ogni valore si ottiene con somme e differenze dei dati, quindi è intero per dati interi.

□

Proposizione 4: i vertici sono soluzioni di base

Ogni vertice $\bar{\boldsymbol{x}}$ *della regione ammissibile del modello* TP *è una soluzione ammissibile di base: le sue celle con* $\bar{x}_{ij} > 0$ *formano una foresta, contenuta in una base* T*, e* $\bar{\boldsymbol{x}}$ *è la soluzione di base di* T*.*

Dimostrazione (1/2)

Sia F l'insieme delle celle con $\bar{x}_{ij} > 0$.

1. Le colonne di F sono linearmente indipendenti, per il primo passo della dimostrazione del Teorema «vertici interi» del Capitolo 4: qui le celle strettamente comprese fra i loro limiti sono quelle di F.

   Un ciclo ha colonne dipendenti (Lemma «colonne di un ciclo» del Capitolo 1), quindi F è una foresta.

Dimostrazione (2/2)

2. La rete è connessa: finché la foresta ha due o più componenti, qualche arco ne collega una a un'altra, e aggiungerlo non crea cicli (Lemma «aggiungere uno spigolo» del Capitolo 1). Si arriva a un albero ricoprente $T \supseteq F$.

3. $\bar{\boldsymbol{x}}$ è nullo fuori da $F \subseteq T$, quindi è la soluzione di base di T, unica per il Teorema «le basi sono alberi ricoprenti».

□

Le basi nella tabella

- Un ciclo della rete bipartita alterna origini e destinazioni
- Nella tabella è un cammino chiuso fra celle dell'insieme, che si muove alternativamente lungo una riga e lungo una colonna
- Un insieme di celle è una base se ha $m + n - 1$ celle e nessun cammino chiuso di questo tipo
- Una soluzione ammissibile di base ha al più $m + n - 1$ quantità positive; ne ha di meno quando è degenere

Esempio: basi dell'istanza di trasporto (1/2)

Nell'esempio dei tre magazzini una base ha
$m + n - 1 = 3 + 4 - 1 = 6$ celle.

- Le celle $(1, 4)$, $(1, 5)$, $(2, 4)$, $(2, 5)$, $(3, 6)$, $(3, 7)$ non sono una base
- Le celle $(1, 4)$, $(2, 4)$, $(2, 5)$, $(1, 5)$ formano il ciclo $1, 4, 2, 5, 1$, e infatti

$$\boldsymbol{A}_{14} - \boldsymbol{A}_{24} + \boldsymbol{A}_{25} - \boldsymbol{A}_{15} = 0$$

per le loro colonne $\boldsymbol{A}_{ij}$

Esempio: basi dell'istanza di trasporto (2/2)

- Le celle $(1,4)$, $(1,5)$, $(1,6)$, $(1,7)$, $(2,7)$, $(3,7)$ formano un albero ricoprente
- Le sue foglie $2,3,4,5,6$ fissano

$$x_{27} = 30, \quad x_{37} = 30, \quad x_{14} = 10, \quad x_{15} = 25, \quad x_{16} = 15$$

- Poi la riga dell'origine 1 dà

$$x_{17} = 30 - 10 - 25 - 15 = -20.$$

- La soluzione di base esiste, ma non è ammissibile

Riepilogo

Rango:

- i vincoli del modello TP hanno rango $m + n - 1$

Basi:

- una base è un albero ricoprente della rete bipartita, con $m + n - 1$ celle
- la sua soluzione di base si trova foglia per foglia, ed è intera per dati interi

Vertici:

- ogni vertice è una soluzione ammissibile di base

Soluzioni ammissibili di base iniziali

Introduzione: Soluzioni ammissibili di base iniziali

- Il metodo dei potenziali parte da una soluzione ammissibile di base: serve un modo per costruirne una
- Riempiamo la tabella una cella alla volta, barrando una riga o una colonna a ogni passo
- Due regole scelgono la cella successiva: la regola dell'angolo nord-ovest e la regola del costo minimo

Linee attive e residui

- Una soluzione ammissibile di base si costruisce riempiendo la tabella una cella alla volta
- Una riga o una colonna è **attiva** finché non viene barrata
- L'offerta residua $\bar{a}_i$ e la domanda residua $\bar{d}_j$ sono ciò che resta da spedire e da ricevere
- Le due regole che seguono differiscono solo nella scelta della cella successiva

Algorithm: Soluzione ammissibile di base iniziale

Input : un'istanza bilanciata con offerte a_i, domande d_j e costi c_{ij}

Output : una base T e la sua soluzione ammissibile di base $\bar{\boldsymbol{x}}$

1 **Passo 0 (inizializzazione).** Tutte le righe e le colonne sono attive; $\bar{a}_i := a_i$, $\bar{d}_j := d_j$, $T := \emptyset$.

2 **Passo 1 (scelta della cella).** Scegli una cella (i, j) con riga e colonna attive:

3 *regola dell'angolo nord-ovest*: la cella con i più piccolo e j più piccolo;

4 *regola del costo minimo*: la cella con c_{ij} più piccolo (a parità, i più piccolo, poi j).

5 **Passo 2 (quantità).** Poni $\bar{x}_{ij} := \min\{\bar{a}_i, \bar{d}_j\}$, aggiungi (i, j) a T, e sottrai $\bar{x}_{ij}$ da $\bar{a}_i$ e da $\bar{d}_j$.

6 **Passo 3 (barratura).** Se $\bar{a}_i = 0$ e un'altra riga è attiva, barra la riga i; altrimenti barra la colonna j.

7 Se nessuna colonna è attiva, `STOP`; altrimenti torna al Passo 1.

Parità: una cella di base con valore 0

- Quando $\bar{a}_i = \bar{d}_j = 0$ nello stesso momento, viene barrata solo la riga
- La colonna resta attiva con domanda residua 0, e una cella successiva di quella colonna riceve la quantità 0
- Questa cella di base con valore 0 mantiene T un albero ricoprente, e la soluzione è degenere

Teorema 2: la soluzione iniziale è di base

Con entrambe le regole, l'algoritmo «Soluzione ammissibile di base iniziale» si ferma dopo $m + n - 1$ *passi, e* T *è una base con soluzione ammissibile di base* $\bar{\mathbf{x}}$.

Dimostrazione (1/3)

1. *Invariante.* Le offerte residue delle righe attive e le domande residue delle colonne attive hanno somme uguali: vale all'inizio, e il Passo 2 sottrae la stessa quantità dalle due parti.

2. *Una linea barrata ha residuo* 0. Una riga barrata ha $\bar{a}_i = 0$. Una colonna viene barrata:
   - quando $\bar{a}_i > 0$, e allora $\bar{x}_{ij} = \bar{d}_j$ e $\bar{d}_j = 0$,
   - oppure quando la riga i è l'unica riga attiva con $\bar{a}_i = 0$, e allora ogni colonna attiva ha residuo 0 per l'invariante.

Dimostrazione (2/3)

3. *Conteggio.* Una riga viene barrata solo se un'altra riga è attiva, quindi una riga r non viene mai barrata, e l'algoritmo si ferma quando le n colonne sono barrate.

   All'arresto le righe attive hanno residuo 0 per l'invariante, e una riga che arriva a 0 senza essere barrata è l'unica riga attiva.

   Quindi i passi sono $(m-1)+n = m+n-1$, una cella ciascuno, tutte diverse, perché una cella si sceglie in due linee attive e una delle due viene poi barrata.

Dimostrazione (3/3)

4. *Ammissibilità.* All'arresto ogni riga e ogni colonna ha residuo 0, quindi $\bar{\boldsymbol{x}} \geq 0$ soddisfa tutti i vincoli.

5. *Albero.* Sia L_t la linea barrata al passo t: nessuna cella successiva sta in L_t.

   Procedendo all'indietro dall'ultimo passo, le celle dei passi $t, t+1, \ldots$ formano un albero: la cella del passo t collega il nuovo nodo L_t alla sua altra linea, ancora attiva, quindi un nodo delle celle successive o la riga r.

   Per $t = 1$ l'albero ricopre tutti gli $m + n$ nodi.

□

Esempio: la regola dell'angolo nord-ovest (1/3)

Sull'esempio dei tre magazzini la regola dell'angolo nord-ovest parte dalla cella $(1, 4)$ e si sposta a destra o in basso. I suoi passi sono:

passo	cella	quantità	barrata
1	$(1, 4)$	10	colonna 4
2	$(1, 5)$	20	riga 1
3	$(2, 5)$	5	colonna 5
4	$(2, 6)$	15	colonna 6
5	$(2, 7)$	10	riga 2
6	$(3, 7)$	30	riga 3, colonna 7

Esempio: la regola dell'angolo nord-ovest (2/3)

La soluzione ammissibile di base, con le celle di base in blu, è

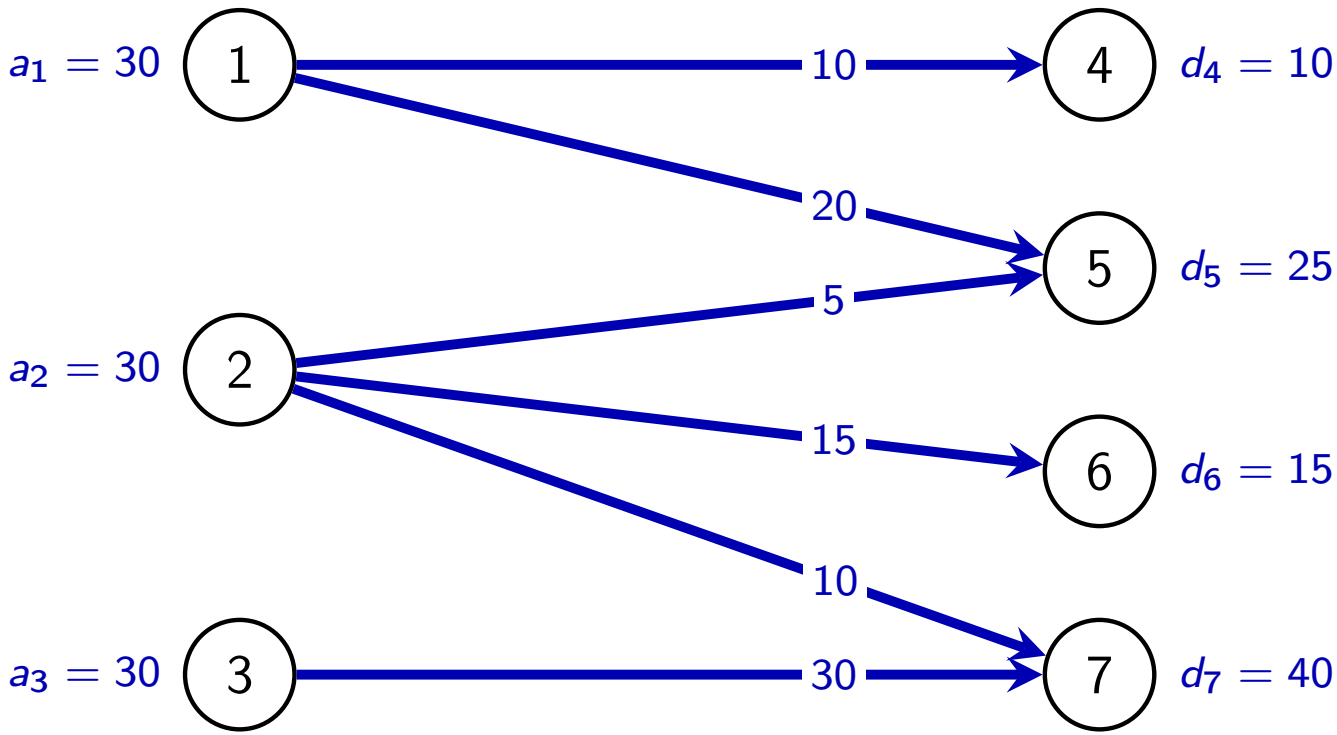
$\bar{x}_{ij}$	4	5	6	7	a_i
1	10	20			30
2		5	15	10	30
3				30	30
d_j	10	25	15	40	

Il suo costo è

$$1 \cdot 10 + 9 \cdot 20 + 4 \cdot 5 + 2 \cdot 15 + 2 \cdot 10 + 6 \cdot 30 = 440.$$

Esempio: la regola dell'angolo nord-ovest (3/3)

Le sei celle di base formano l'albero ricoprente qui sotto, con le quantità sugli archi.



La regola dell'angolo nord-ovest ignora i costi

- La regola dell'angolo nord-ovest guarda solo la posizione delle celle, mai il loro costo
- Su questa istanza spedisce 20 unità sulla cella $(1, 5)$, di costo 9, e 30 unità su $(3, 7)$, di costo 6
- La regola del costo minimo, invece, guarda i costi

Esempio: la regola del costo minimo (1/3)

Sull'esempio dei tre magazzini la regola del costo minimo sceglie, a ogni passo, la cella più economica fra le righe e le colonne attive:

passo	cella	quantità	barrata
1	$(1,4)$	10	colonna 4
2	$(3,5)$	25	colonna 5
3	$(1,6)$	15	colonna 6
4	$(2,7)$	30	riga 2
5	$(1,7)$	5	riga 1
6	$(3,7)$	5	riga 3, colonna 7

Esempio: la regola del costo minimo (2/3)

- Al passo 1 le celle $(1, 4)$ e $(3, 5)$ costano entrambe 1, e $(1, 4)$ viene prima, perché la sua riga è più piccola
- Al passo 3 le celle $(1, 6)$, $(2, 6)$, $(2, 7)$ costano 2, e $(1, 6)$ viene prima
- Al passo 4 è attiva solo la colonna 7: prende tutta l'offerta della riga 2, poi 5 unità dalla riga 1 e 5 dalla riga 3

Esempio: la regola del costo minimo (3/3)

La soluzione ammissibile di base è

$\bar{x}_{ij}$	4	5	6	7	a_i
1	10		15	5	30
2				30	30
3		25		5	30
d_j	10	25	15	40	

con costo

$$1 \cdot 10 + 2 \cdot 15 + 3 \cdot 5 + 2 \cdot 30 + 1 \cdot 25 + 6 \cdot 5 = 170.$$

Riepilogo

Costruire una soluzione ammissibile di base:

- una cella per passo e una linea barrata per passo
- il risultato ha $m + n - 1$ celle di base, ed è una base

Le due regole:

- la regola dell'angolo nord-ovest ignora i costi
- la regola del costo minimo di solito parte molto più vicino all'ottimo: 170 invece di 440 nell'esempio

Il metodo dei potenziali

Introduzione: Il metodo dei potenziali

- Da una soluzione ammissibile di base vogliamo arrivare a una soluzione ottima
- I potenziali della base danno i costi ridotti delle celle: quando sono tutti non negativi la soluzione è ottima, e i potenziali sono una soluzione duale ottima
- Altrimenti una cella con costo ridotto negativo entra in base, e le quantità si spostano lungo il suo ciclo: è l'algoritmo del simplesso, scritto nella tabella

I potenziali come variabili duali

- Scriviamo i vincoli di domanda come

$$-\sum_{i \in P} x_{ij} = -d_j,$$

  i vincoli di conservazione delle destinazioni

- Usiamo una variabile duale libera π_v, il potenziale del nodo v, per il vincolo di ogni nodo $v \in P \cup D$, come nel Capitolo 4

Il modello $D(\mathrm{TP})$

Con una variabile duale libera π_v per il vincolo di ogni nodo $v \in P \cup D$, il modello duale del modello TP è:

$$\max \quad \sum_{i \in P} a_i \, \pi_i - \sum_{j \in D} d_j \, \pi_j$$

$$\text{soggetto a} \quad \pi_i - \pi_j \leq c_{ij}, \quad \forall i \in P, \ \forall j \in D,$$

$$\pi_v \in \mathbb{R}, \quad \forall v \in P \cup D.$$

Questo modello LP è detto modello $D(\mathrm{TP})$.

Costi ridotti e normalizzazione

- Con il costo ridotto del Capitolo 4

$$\bar{c}_{ij} = c_{ij} - \pi_i + \pi_j,$$

il vincolo duale $\pi_i - \pi_j \leq c_{ij}$ diventa $\bar{c}_{ij} \geq 0$

- Aggiungere la stessa costante a tutti i potenziali non cambia i costi ridotti né, in un'istanza bilanciata, l'obiettivo duale

- Normalizziamo con $\pi_{m+n} = 0$, il potenziale dell'ultima destinazione

Costo e costo ridotto

Per ogni $\boldsymbol{x}$ che soddisfa i vincoli di offerta e di domanda, i potenziali si sommano riga per riga e colonna per colonna:

$$\sum_{i \in P} \sum_{j \in D} (\pi_i - \pi_j)\, x_{ij} = \sum_{i \in P} a_i\, \pi_i - \sum_{j \in D} d_j\, \pi_j,$$

e poiché $c_{ij} = \bar{c}_{ij} + \pi_i - \pi_j$,

$$\sum_{i \in P} \sum_{j \in D} c_{ij}\, x_{ij} = \sum_{i \in P} \sum_{j \in D} \bar{c}_{ij}\, x_{ij} + \sum_{i \in P} a_i\, \pi_i - \sum_{j \in D} d_j\, \pi_j.$$

È il Lemma «costo e costo ridotto» del Capitolo 4 per la rete bipartita.

Definizione 3: potenziali di una base

*I **potenziali di una base** T sono i valori π_v, $v \in P \cup D$, tali che*

$$\bar{c}_{ij} = c_{ij} - \pi_i + \pi_j = 0, \quad \forall (i,j) \in T,$$

$$\pi_{m+n} = 0.$$

- I potenziali di una base esistono e sono unici

Calcolare i potenziali di una base

- Si parte da $\pi_{m+n} = 0$
- Una cella di base (i, j) con un estremo noto dà l'altro:

$$\pi_i = \pi_j + c_{ij} \qquad \text{oppure} \qquad \pi_j = \pi_i - c_{ij}$$

- Ogni nodo si raggiunge lungo il suo unico cammino nell'albero a partire da $m + n$, che ne fissa il valore

Teorema 3: criterio di ottimalità

Sia $\bar{\mathbf{x}}$ *la soluzione ammissibile di base di una base* T*, con potenziali* $\boldsymbol{\pi}$*. Se* $\bar{c}_{ij} \geq 0$ *per ogni cella, allora* $\bar{\mathbf{x}}$ *è ottima per il modello* TP*,* $\boldsymbol{\pi}$ *è ottima per il modello* $D(\mathrm{TP})$*, e*

$$\sum_{i \in P} \sum_{j \in D} c_{ij}\, \bar{x}_{ij} = \sum_{i \in P} a_i\, \pi_i - \sum_{j \in D} d_j\, \pi_j.$$

Se inoltre $\bar{c}_{ij} > 0$ *per ogni cella fuori da* T*, allora* $\bar{\mathbf{x}}$ *è l'unica soluzione ottima.*

Dimostrazione (1/2)

Usiamo l'identità «costo e costo ridotto». Sia

$$z_\pi = \sum_{i \in P} a_i \pi_i - \sum_{j \in D} d_j \pi_j.$$

1. Per $\bar{\boldsymbol{x}}$: $\bar{c}_{ij} = 0$ su T e $\bar{x}_{ij} = 0$ fuori da T, quindi $\sum \bar{c}_{ij} \bar{x}_{ij} = 0$ e il costo di $\bar{\boldsymbol{x}}$ è z_π.

2. Per ogni $\boldsymbol{x}$ ammissibile: $\bar{c} \geq 0$ e $\boldsymbol{x} \geq 0$, quindi il suo costo è almeno z_π; dunque $\bar{\boldsymbol{x}}$ è ottima.

Dimostrazione (2/2)

I potenziali soddisfano i vincoli duali $\pi_i - \pi_j \leq c_{ij}$, e il loro valore z_π è uguale al costo di una soluzione primale: per la dualità debole sono ottimi.

3. Se $\bar{c}_{ij} > 0$ fuori da T, una $\boldsymbol{x}$ ottima ha $\sum \bar{c}_{ij} x_{ij} = 0$, quindi $x_{ij} = 0$ fuori da T: è la soluzione di base di T, che è $\bar{\boldsymbol{x}}$ per il Teorema «le basi sono alberi ricoprenti».

□

Il ciclo di una cella e il pivot

- Quando qualche cella ha $\bar{c}_e < 0$, il criterio non vale, e cerchiamo di migliorare la soluzione
- La cella e chiude un ciclo con l'albero della base: spostare quantità lungo questo ciclo mantiene bilanciate tutte le righe e le colonne
- Il lemma che segue dà i segni delle celle del ciclo, e il teorema successivo il pivot

Lemma 1: il ciclo di una cella non di base (1/2)

Siano T una base ed $e = (k, l) \notin T$. Le celle $T \cup \{e\}$ contengono esattamente un ciclo C, formato da e e dal cammino da l a k nell'albero.

Lemma: il ciclo di una cella non di base (2/2)

Scriviamo C partendo da e e muovendoci prima lungo la riga k, e diamo il segno $+$ alle celle nelle posizioni dispari (e compresa) e $-$ alle altre.

Allora ogni riga e ogni colonna che incontra C contiene esattamente una cella di C con segno $+$ e una con segno $-$.

Dimostrazione

- Il ciclo è il ciclo fondamentale dell'arco (k, l) (Proposizione «ciclo fondamentale» del Capitolo 1)
- Visita i nodi $k, j_1, i_1, j_2, \ldots, l$ alternativamente in P e in D, quindi ha un numero pari di celle, e due celle consecutive hanno un nodo in comune: le prime due la riga k, le due successive una colonna, e così via
- Ogni nodo di C sta in esattamente due celle di C, consecutive lungo C, compresa la coppia formata dall'ultima e dalla prima: le loro posizioni hanno parità diversa, quindi segni diversi

□

Teorema 4: il pivot (1/2)

Siano $\bar{\boldsymbol{x}}$ *la soluzione ammissibile di base di una base* T *con potenziali* $\boldsymbol{\pi}$, $e \notin T$ *con* $\bar{c}_e < 0$, C^+ *e* C^- *le celle del suo ciclo con segno* $+$ *e* $-$, *e*

$$\theta = \min\{\bar{x}_f : f \in C^-\},$$

raggiunto in una cella $f^* \in C^-$. *Per* $0 \leq t \leq \theta$, *sia* $\boldsymbol{x}(t)$ *il vettore* $\bar{\boldsymbol{x}}$ *con* t *aggiunto sulle celle di* C^+ *e sottratto sulle celle di* C^-.

Teorema: il pivot (2/2)

(a) $\boldsymbol{x}(t)$ *è ammissibile, e il suo costo è il costo di* $\bar{\boldsymbol{x}}$ *più* $t\,\bar{c}_e$.

(b) $T' = T \cup \{e\} \setminus \{f^*\}$ *è una base, e* $\boldsymbol{x}(\theta)$ *è la sua soluzione ammissibile di base.*

Dimostrazione (1/2)

1. *Ammissibilità.* Per il Lemma «il ciclo di una cella non di base» ogni riga e colonna di C guadagna t su una cella e perde t su un'altra, quindi tutte le somme restano invariate. Le celle di C^- restano $\bar{x}_f - t \geq 0$ perché $t \leq \theta$.

2. *Costo.* La variazione di costo è $t\big(\sum_{C^+} c_{ij} - \sum_{C^-} c_{ij}\big)$. Ogni nodo di C sta in una cella di C^+ e in una di C^-, quindi

$$\sum_{C^+}(\pi_i - \pi_j) = \sum_{C^-}(\pi_i - \pi_j),$$

e la variazione è $t\big(\sum_{C^+} \bar{c}_{ij} - \sum_{C^-} \bar{c}_{ij}\big) = t\,\bar{c}_e$: le altre celle di C sono di base, con $\bar{c}_{ij} = 0$.

Dimostrazione (2/2)

3. *Nuova base.* T' ha $m+n-1$ celle, e non ha cicli: l'unico ciclo di $T \cup \{e\}$ è C, che contiene f^*.

   Un insieme aciclico di $m+n-1$ archi su $m+n$ nodi è un albero ricoprente (Teorema «caratterizzazioni di un albero» del Capitolo 1).

   $\boldsymbol{x}(\theta)$ è nullo fuori da T', perché $x_{f^*}(\theta) = 0$, quindi è la soluzione di base di T'.

□

Algorithm: Il metodo dei potenziali

Input : un'istanza bilanciata con offerte a_i, domande d_j e costi c_{ij}

Output : una soluzione ammissibile di base ottima $\tilde{\boldsymbol{x}}$ e potenziali ottimi $\tilde{\boldsymbol{\pi}}$

1. **Passo 0 (base iniziale).** Costruisci una base T e la sua soluzione ammissibile di base $\bar{\boldsymbol{x}}$ con la regola dell'angolo nord-ovest o con la regola del costo minimo.
2. **Passo 1 (potenziali).** Calcola i potenziali di T e i costi ridotti $\bar{c}_{ij} = c_{ij} - \pi_i + \pi_j$ delle celle fuori da T.
3. **Passo 2 (test di ottimalità).** Se $\bar{c}_{ij} \geq 0$ per ogni cella, `STOP`: $\bar{\boldsymbol{x}}$ è ottima.
4. **Passo 3 (cella entrante).** Scegli una cella e con $\bar{c}_e < 0$.
5. **Passo 4 (ciclo).** Trova il ciclo di e con i suoi segni, $\theta := \min\{\bar{x}_f : f \in C^-\}$, e una cella uscente $f^* \in C^-$ con $\bar{x}_{f^*} = \theta$.
6. **Passo 5 (pivot).** Aggiungi θ su C^+, sottrai θ su C^-, poni $T := T \cup \{e\} \setminus \{f^*\}$, e torna al Passo 1.

Regole degli esempi e nomi

- **Regole degli esempi.** Al Passo 3 la cella entrante ha il costo ridotto più negativo (a parità, i più piccolo, poi j). Al Passo 4 il ciclo si scrive dalla cella entrante muovendosi prima lungo la sua riga, e la cella uscente è la prima cella di C^- con $\bar{x}_f = \theta$ in ordine riga–colonna.

- Il metodo è noto anche come **metodo u-v** o metodo MODI (*modified distribution*), e il ciclo come cammino **stepping stone**: le quantità si spostano lungo il ciclo come sulle pietre per attraversare un fiume.

Teorema 5: terminazione

Se $\theta > 0$ a ogni iterazione, il metodo dei potenziali si ferma dopo un numero finito di iterazioni con una soluzione ammissibile di base ottima e potenziali ottimi.

Dimostrazione

- Per il Teorema «il pivot» ogni iterazione con $\theta > 0$ passa a una nuova base e diminuisce il costo di $\theta\,|\bar{c}_e| > 0$
- Una base determina la sua soluzione di base (Teorema «le basi sono alberi ricoprenti»), quindi una base visitata due volte avrebbe due costi diversi: nessuna base si ripete
- Gli insiemi di $m + n - 1$ celle sono in numero finito, quindi il metodo si ferma, al Passo 2, e si applica il Teorema «criterio di ottimalità»

□

Pivot degeneri e algoritmo del simplesso

- Quando $\theta = 0$ il pivot è **degenere**: cambia la base, ma non la soluzione né il suo costo
- Il metodo dei potenziali è l'algoritmo del simplesso sul modello TP: le celle di base sono le variabili di base, i $\bar{c}_{ij}$ sono i loro costi ridotti, e il ciclo descrive come cambiano le variabili di base quando x_e cresce
- Si applica quindi la regola anticiclo di Bland del Capitolo 7 «Condizioni di ottimalità e regole di pivot» del modulo Programmazione lineare: con la prima cella con $\bar{c}_{ij} < 0$ in ordine riga–colonna come cella entrante, e la regola di uscita vista sopra, nessuna base si ripete, anche con pivot degeneri

Osservazione 1: La convenzione u-v

Molti libri di testo scrivono il metodo u-v con $u_i + v_j = c_{ij}$ *sulle celle di base e* $\bar{c}_{ij} = c_{ij} - u_i - v_j$.

Sono i potenziali di questo capitolo con $u_i = \pi_i$ *per le origini e* $v_j = -\pi_j$ *per le destinazioni: i costi ridotti sono gli stessi, e i potenziali delle destinazioni cambiano segno.*

Il metodo dei potenziali sull'istanza di trasporto

- Partiamo dalla soluzione dell'angolo nord-ovest dell'esempio dei tre magazzini, di costo 440
- In ogni tabella le celle di base (blu) contengono la quantità $\bar{x}_{ij}$ in grassetto, e le altre celle il costo ridotto $\bar{c}_{ij}$
- L'ultima colonna e l'ultima riga contengono i potenziali
- La cella entrante è rossa, e le celle del ciclo sono segnate con $\oplus$ e $\ominus$

Esempio: il metodo dei potenziali dall'angolo nord-ovest (1/13)

Nella prima iterazione i potenziali della base nord-ovest si ricavano da $\pi_7 = 0$, lungo l'albero:

$$\pi_2 = \pi_7 + c_{27} = 2, \quad \pi_3 = \pi_7 + c_{37} = 6, \quad \pi_5 = \pi_2 - c_{25} = -2,$$

$$\pi_6 = \pi_2 - c_{26} = 0, \quad \pi_1 = \pi_5 + c_{15} = 7, \quad \pi_4 = \pi_1 - c_{14} = 6.$$

Per esempio,

$$\bar{c}_{35} = c_{35} - \pi_3 + \pi_5 = 1 - 6 - 2 = -7.$$

Esempio: il metodo dei potenziali dall'angolo nord-ovest (2/13)

Iterazione 1

	4	5	6	7	π_i
1	**10**	**20**	-5	-4	7
2	7	**5** $\ominus$	**15**	**10** $\oplus$	2
3	9	-7 $\oplus$	-2	**30** $\ominus$	6
π_j	6	-2	0	0	

- I costi ridotti negativi sono $\bar{c}_{35} = -7$, $\bar{c}_{16} = -5$, $\bar{c}_{17} = -4$, $\bar{c}_{36} = -2$; la cella $(3, 5)$ entra in base.

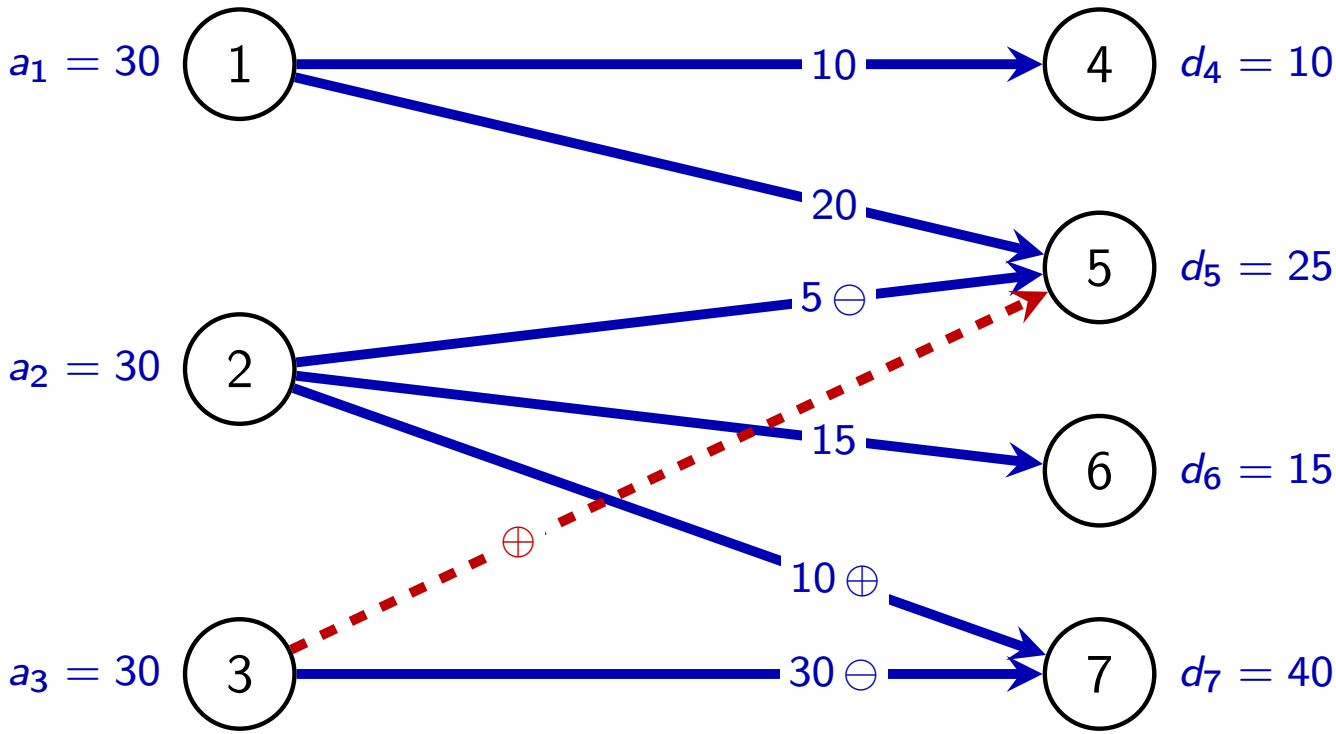
Esempio: il metodo dei potenziali dall'angolo nord-ovest (3/13)

Iterazione 1:

- Il ciclo è $(3,5)^+,\ (3,7)^-,\ (2,7)^+,\ (2,5)^-$.
- Le celle con segno $-$ hanno $\bar{x}_{37} = 30$ e $\bar{x}_{25} = 5$: $\theta = 5$, e la cella $(2,5)$ esce dalla base.
- Il costo passa da 440 a $440 + 5 \cdot (-7) = 405$.

Esempio: il metodo dei potenziali dall'angolo nord-ovest (4/13)

Nella rete, l'arco entrante $(3, 5)$ (rosso, tratteggiato) chiude il ciclo $3, 5, 2, 7, 3$ con l'albero.



Esempio: il metodo dei potenziali dall'angolo nord-ovest (5/13)

Nella seconda iterazione il ciclo ha sei celle e visita tutte e tre le righe.

Iterazione 2

	4	5	6	7	π_i
1	**10**	**20** $\ominus$	-12 $\oplus$	-11	14
2	14	7	**15** $\ominus$	**15** $\oplus$	2
3	16	**5** $\oplus$	-2	**25** $\ominus$	6
π_j	13	5	0	0	

- I costi ridotti negativi sono $\bar{c}_{16} = -12$, $\bar{c}_{17} = -11$, $\bar{c}_{36} = -2$; la cella $(1, 6)$ entra in base.

Esempio: il metodo dei potenziali dall'angolo nord-ovest (6/13)

Iterazione 2:

- Il ciclo è $(1,6)^+,\ (1,5)^-,\ (3,5)^+,\ (3,7)^-,\ (2,7)^+,\ (2,6)^-$.
- Le celle con segno $-$ hanno $\bar{x}_{15} = 20$, $\bar{x}_{37} = 25$ e $\bar{x}_{26} = 15$: $\theta = 15$, e la cella $(2,6)$ esce dalla base.
- Il costo passa da 405 a $405 + 15 \cdot (-12) = 225$.

Esempio: il metodo dei potenziali dall'angolo nord-ovest (7/13)

Iterazione 3

	4	5	6	7	π_i
1	**10**	**5** $\ominus$	**15**	-11 $\oplus$	14
2	14	7	12	**30**	2
3	16	**20** $\oplus$	10	**10** $\ominus$	6
π_j	13	5	12	0	

- L'unico costo ridotto negativo è $\bar{c}_{17} = -11$: la cella $(1, 7)$ entra in base.

Esempio: il metodo dei potenziali dall'angolo nord-ovest (8/13)

Iterazione 3:

- Il ciclo è $(1,7)^+,\ (1,5)^-,\ (3,5)^+,\ (3,7)^-$.
- Le celle con segno $-$ hanno $\bar{x}_{15} = 5$ e $\bar{x}_{37} = 10$: $\theta = 5$, e la cella $(1,5)$ esce dalla base.
- Il costo passa da 225 a $225 + 5 \cdot (-11) = 170$.

Esempio: il metodo dei potenziali dall'angolo nord-ovest (9/13)

Iterazione 4

	4	5	6	7	π_i
1	**10**	11	**15** $\ominus$	**5** $\oplus$	3
2	3	7	1	**30**	2
3	5	**25**	-1 $\oplus$	**5** $\ominus$	6
π_j	2	5	1	0	

- L'unico costo ridotto negativo è $\bar{c}_{36} = -1$: la cella $(3, 6)$ entra in base.

Esempio: il metodo dei potenziali dall'angolo nord-ovest (10/13)

Iterazione 4:

- Il ciclo è $(3,6)^+,\ (3,7)^-,\ (1,7)^+,\ (1,6)^-$.
- Le celle con segno $-$ hanno $\bar{x}_{37} = 5$ e $\bar{x}_{16} = 15$: $\theta = 5$, e la cella $(3,7)$ esce dalla base.
- Il costo passa da 170 a $170 + 5 \cdot (-1) = 165$.

Esempio: il metodo dei potenziali dall'angolo nord-ovest (11/13)

Iterazione 5

	4	5	6	7	π_i
1	**10**	10	**10**	**10**	3
2	3	6	1	**30**	2
3	6	**25**	**5**	1	5
π_j	2	4	1	0	

- Tutti i costi ridotti sono non negativi: la soluzione ammissibile di base è ottima, con costo $z = 165$.

Esempio: il metodo dei potenziali dall'angolo nord-ovest (12/13)

- La soluzione ottima è

$$\tilde{x}_{14} = 10, \quad \tilde{x}_{16} = 10, \quad \tilde{x}_{17} = 10,$$

$$\tilde{x}_{27} = 30, \quad \tilde{x}_{35} = 25, \quad \tilde{x}_{36} = 5, \qquad z(\mathrm{TP}) = 165.$$

- Ogni cella fuori dalla base ha $\bar{c}_{ij} > 0$, quindi per il Teorema «criterio di ottimalità» l'ottimo è unico.

Esempio: il metodo dei potenziali dall'angolo nord-ovest (13/13)

- I potenziali $\tilde{\boldsymbol{\pi}} = (3, 2, 5, 2, 4, 1, 0)$ sono ottimi per il modello $D(\mathrm{TP})$, con valore

  $$30\,(3+2+5) - (10\cdot 2 + 25\cdot 4 + 15\cdot 1 + 40\cdot 0) = 300 - 135 = 165.$$

  Gurobi restituisce la stessa soluzione ottima, e i suoi valori duali coincidono con $\tilde{\boldsymbol{\pi}}$.

- Dalla soluzione del costo minimo il metodo richiede un solo pivot: quella soluzione ha la base dell'iterazione 4, e la cella entrante $(3, 6)$, con $\theta = 5$, dà l'ottimo.

Degenerazione

- Quando una cella di base ha valore 0, un pivot può avere $\theta = 0$
- Eseguiamo ora il metodo dei potenziali su un'istanza in cui la regola dell'angolo nord-ovest dà una soluzione degenere
- Vediamo un pivot degenere, e perché la cella di base con valore 0 va tenuta

Esempio: un pivot degenere (1/6)

Consideriamo l'istanza

c_{ij}	4	5	6	a_i
1	1	4	7	30
2	6	2	5	25
3	3	9	5	30
d_j	30	30	25	85

- Nella regola dell'angolo nord-ovest la prima cella esaurisce insieme la riga 1 e la colonna 4.

Esempio: un pivot degenere (2/6)

- Viene barrata solo la riga, e la cella $(2, 4)$ entra in T con quantità 0: la soluzione, di costo 250, è degenere.

passo	cella	quantità	barrata
1	$(1, 4)$	30	riga 1
2	$(2, 4)$	0	colonna 4
3	$(2, 5)$	25	riga 2
4	$(3, 5)$	5	colonna 5
5	$(3, 6)$	25	riga 3, colonna 6

Esempio: un pivot degenere (3/6)

Iterazione 1

	4	5	6	π_i
1	**30**	7	14	-7
2	**0** $\ominus$	**25** $\oplus$	7	-2
3	-10 $\oplus$	**5** $\ominus$	**25**	5
π_j	-8	-4	0	

- L'unico costo ridotto negativo è $\bar{c}_{34} = -10$: la cella $(3,4)$ entra in base.
- Il ciclo è $(3,4)^+,\ (3,5)^-,\ (2,5)^+,\ (2,4)^-$.

Esempio: un pivot degenere (4/6)

Iterazione 1:

- Le celle con segno $-$ hanno $\bar{x}_{35} = 5$ e $\bar{x}_{24} = 0$: $\theta = 0$, e la cella $(2, 4)$ esce dalla base.

- Il pivot è degenere: la soluzione e il suo costo $z = 250$ non cambiano, cambia solo la base.

Nella seconda iterazione la base è cambiata, quindi sono cambiati anche i potenziali: ora $\bar{c}_{15} < 0$.

Esempio: un pivot degenere (5/6)

Iterazione 2

	4	5	6	π_i
1	**30** $\ominus$	-3 $\oplus$	4	3
2	10	**25**	7	-2
3	**0** $\oplus$	**5** $\ominus$	**25**	5
π_j	2	-4	0	

- L'unico costo ridotto negativo è $\bar{c}_{15} = -3$: la cella $(1,5)$ entra in base.
- Il ciclo è $(1,5)^+,\ (1,4)^-,\ (3,4)^+,\ (3,5)^-$.

Esempio: un pivot degenere (6/6)

Iterazione 2:

- Le celle con segno $-$ hanno $\bar{x}_{14} = 30$ e $\bar{x}_{35} = 5$: $\theta = 5$, e la cella $(3, 5)$ esce dalla base.
- Il costo passa da 250 a $250 + 5 \cdot (-3) = 235$.

Iterazione 3

	4	5	6	π_i
1	**25**	**5**	4	3
2	7	**25**	4	1
3	**5**	3	**25**	5
π_j	2	-1	0	

Tutti i costi ridotti sono non negativi: la soluzione ammissibile di base è ottima, con costo $z = 235$. È unica, e Gurobi lo conferma.

Che cosa mostra l'esempio degenere

- La cella di base con valore 0 va tenuta in T: senza di essa, le altre quattro celle non ricoprirebbero i sei nodi, e i potenziali non si potrebbero calcolare
- Il pivot degenere dell'iterazione 1 non abbassa il costo, ma porta a una base i cui potenziali rivelano una cella migliorante

Riepilogo (1/2)

Potenziali:

- $\bar{c}_{ij} = 0$ sulle celle di base e $\pi_{m+n} = 0$
- sono le variabili duali del modello $D(\mathrm{TP})$

Ottimalità:

- il metodo si ferma quando ogni costo ridotto è non negativo
- allora i potenziali sono una soluzione duale ottima, con lo stesso valore

Riepilogo (2/2)

Pivot:

- altrimenti θ unità si spostano lungo il ciclo di una cella entrante
- il costo diminuisce di $\theta\,|\bar{c}_e|$

Degenerazione:

- con $\theta = 0$ cambia solo la base
- le celle di base con valore 0 restano in base

Il problema di assegnamento

Introduzione: Il problema di assegnamento

- n lavoratori vanno assegnati a n lavori, un lavoro ciascuno, a costo totale minimo
- È un problema di trasporto con offerte e domande unitarie, ma il metodo dei potenziali vi fa molti pivot degeneri
- Dimostriamo il teorema di König ed Egerváry su zeri e linee, e lo usiamo nell'algoritmo ungherese, che lavora direttamente sulla matrice dei costi

Definizione 4: problema di assegnamento

Ci sono n lavoratori, le origini $P = \{1, 2, \ldots, n\}$, e n lavori, le destinazioni $D = \{n+1, n+2, \ldots, 2n\}$; c_{ij} è il costo di assegnare il lavoratore i al lavoro j.

*Un **assegnamento** dà a ogni lavoratore esattamente un lavoro e ogni lavoro a esattamente un lavoratore.*

*Il **problema di assegnamento** chiede un assegnamento di costo totale minimo.*

- Usiamo una variabile binaria x_{ij} per ogni coppia $i \in P$, $j \in D$, uguale a 1 se il lavoratore i riceve il lavoro j.

Il modello AP

Con queste variabili, un modello di Programmazione Lineare Intera (ILP) per il problema di assegnamento è il seguente:

$$\text{min} \quad \sum_{i \in P} \sum_{j \in D} c_{ij}\, x_{ij}$$

$$\text{soggetto a} \quad \sum_{j \in D} x_{ij} = 1, \qquad \forall i \in P,$$

$$\sum_{i \in P} x_{ij} = 1, \qquad \forall j \in D,$$

$$x_{ij} \in \{0, 1\}, \quad \forall i \in P,\ \forall j \in D.$$

Questo modello è detto modello AP.

Il modello AP: funzione obiettivo e vincoli

Descrizione della funzione obiettivo e dei vincoli del modello AP:

- la funzione obiettivo minimizza il costo totale dell'assegnamento;
- i primi vincoli danno a ogni lavoratore esattamente un lavoro;
- i secondi vincoli danno ogni lavoro a esattamente un lavoratore;
- infine, gli ultimi vincoli definiscono le variabili binarie del modello.

- Una matrice $n \times n$ con elementi in $\{0, 1\}$ ed esattamente un 1 in ogni riga e in ogni colonna è una **matrice di permutazione**
- Le soluzioni ammissibili del modello AP sono le matrici di permutazione: ci sono $n!$ assegnamenti

Proposizione 5: il rilassamento LP del modello di assegnamento

Il rilassamento LP del modello AP*, con* $x_{ij} \geq 0$ *al posto di* $x_{ij} \in \{0, 1\}$*, è il modello* TP *con* $a_i = 1$ *e* $d_j = 1$*. I suoi vertici sono matrici di permutazione, e ha una soluzione ottima che è un assegnamento.*

Dimostrazione

- I vincoli del modello AP sono i vincoli di offerta e di domanda del modello TP con offerte e domande unitarie, e l'istanza è bilanciata, con totale n
- Per la Proposizione «ammissibilità e integralità» ha una soluzione ottima in un vertice, e ogni vertice è intero per il Teorema «vertici interi» del Capitolo 4
- Una soluzione ammissibile intera ha $0 \leq x_{ij} \leq 1$, perché x_{ij} è al più la somma della sua riga: è una matrice di permutazione

□

Esempio: quattro lavoratori e quattro lavori

Quattro lavoratori $1, 2, 3, 4$ e quattro lavori $5, 6, 7, 8$ hanno i costi

c_{ij}	5	6	7	8
1	2	4	3	5
2	4	8	1	1
3	2	6	9	6
4	1	9	8	6

- I $4! = 24$ assegnamenti, enumerati, hanno un unico minimo: il lavoratore 1 riceve il lavoro 7, il lavoratore 2 il lavoro 8, il lavoratore 3 il lavoro 6, il lavoratore 4 il lavoro 5, con costo $3 + 1 + 6 + 1 = 11$.
- Gurobi dà lo stesso assegnamento come soluzione ottima del rilassamento LP.

Degenerazione del metodo dei potenziali

- Una base del rilassamento LP ha $2n - 1$ celle, ma un assegnamento ha solo n celle con $x_{ij} = 1$
- Ogni soluzione ammissibile di base ha $n - 1$ celle di base con valore 0, ed è degenere
- Il metodo dei potenziali risolve il rilassamento LP, ma molti suoi pivot hanno $\theta = 0$

Esempio: la base dell'angolo nord-ovest dell'istanza di assegnamento

Sull'esempio dei quattro lavoratori la regola dell'angolo nord-ovest dà la base qui sotto, con tre zeri di base:

$\bar{x}_{ij}$	5	6	7	8	a_i
1	1				1
2	0	1			1
3		0	1		1
4			0	1	1
d_j	1	1	1	1	

Da questa base il metodo dei potenziali arriva all'ottimo dopo 6 pivot, 4 dei quali degeneri.

Riduzioni e potenziali

- L'algoritmo ungherese lavora sulla matrice dei costi: sottrae costanti dalle righe e dalle colonne
- Queste sottrazioni sono potenziali: la matrice ridotta ha elementi $\bar{c}_{ij} = c_{ij} - \pi_i + \pi_j$
- Il lemma che segue mostra che non cambiano gli assegnamenti ottimi

Lemma 2: costi ridotti di un assegnamento

Dati i potenziali $\boldsymbol{\pi}$*, sia* $\bar{c}_{ij} = c_{ij} - \pi_i + \pi_j$*. Per ogni* $\boldsymbol{x}$ *che soddisfa i vincoli del modello* AP *sulle righe e sulle colonne,*

$$\sum_{i \in P} \sum_{j \in D} c_{ij}\, x_{ij} = \sum_{i \in P} \sum_{j \in D} \bar{c}_{ij}\, x_{ij} + \sum_{i \in P} \pi_i - \sum_{j \in D} \pi_j.$$

Quindi sottrarre una costante da una riga o da una colonna della matrice dei costi cambia il costo di ogni assegnamento della stessa costante, e gli assegnamenti ottimi non cambiano.

Dimostrazione

- È l'identità «costo e costo ridotto» del metodo dei potenziali con $a_i = d_j = 1$
- Sottrarre r dalla riga i dà i costi ridotti di $\pi_i = r$ con gli altri potenziali nulli
- Sottrarre s dalla colonna j dà quelli di $\pi_j = -s$

□

Corollario 1: gli assegnamenti a costo ridotto nullo sono ottimi

Se $\bar{c}_{ij} \geq 0$ *per ogni cella e un assegnamento usa solo celle con* $\bar{c}_{ij} = 0$*, allora è ottimo, anche per il rilassamento LP, il suo costo è* $\sum_{i \in P} \pi_i - \sum_{j \in D} \pi_j$*, e* $\boldsymbol{\pi}$ *è ottima per il duale del rilassamento LP, il modello* $D(\mathrm{TP})$ *con* $a_i = d_j = 1$*.*

Dimostrazione

- Per il Lemma «costi ridotti di un assegnamento», l'assegnamento costa $\sum_{i \in P} \pi_i - \sum_{j \in D} \pi_j$
- Ogni $\boldsymbol{x} \geq 0$ che soddisfa i vincoli sulle righe e sulle colonne costa almeno altrettanto, perché $\bar{c} \geq 0$
- I potenziali soddisfano $\pi_i - \pi_j \leq c_{ij}$ con lo stesso valore, quindi sono ottimi per la dualità debole

□

Definizione 5: zeri indipendenti e coperture

Sia $\bar{c}$ una matrice $n \times n$ con elementi non negativi.

*Un insieme di elementi nulli è **indipendente** se due di essi non stanno mai nella stessa riga o nella stessa colonna.*

*Un insieme di **linee**, righe e colonne, **copre** gli zeri se ogni elemento nullo sta in almeno una delle linee.*

Zeri, matching e coperture

- n zeri indipendenti formano un assegnamento di costo ridotto nullo, ottimo per il Corollario «gli assegnamenti a costo ridotto nullo sono ottimi»

- Nella rete bipartita i cui archi sono le celle con $\bar{c}_{ij} = 0$:
  - un insieme di zeri indipendenti è un matching, cioè un insieme di archi senza estremi in comune, come nel Capitolo 5 «Flusso massimo»,
  - una copertura è un insieme di nodi che tocca ogni arco

Teorema 6: König ed Egerváry

Il numero massimo di zeri indipendenti di una matrice è uguale al numero minimo di linee che coprono tutti i suoi zeri.

Dimostrazione (1/3)

Sia M un insieme massimo di zeri indipendenti. Una riga o una colonna è **accoppiata** se contiene uno zero di M.

1. *Linee $\geq$ zeri.* Gli zeri di M stanno in righe e colonne diverse, quindi una linea ne contiene al più uno, e una copertura richiede almeno $|M|$ linee.

2. *Cammini alternanti.* Un **cammino alternante** parte da una riga non accoppiata, va da una riga a una colonna attraverso uno zero non in M, e da una colonna a una riga attraverso lo zero di M di quella colonna. Siano R e K le righe e le colonne raggiunte da cammini alternanti, comprese le righe non accoppiate.

Dimostrazione (2/3)

3. *Ogni colonna di K è accoppiata.* Altrimenti un cammino alternante termina in una colonna non accoppiata. Lungo di esso, gli zeri non in M sono uno in più degli zeri di M: scambiandoli si ottiene un insieme indipendente con $|M| + 1$ zeri, contro la massimalità di M.

4. *Le linee.* Sia L l'insieme delle righe non in R e delle colonne in K. Uno zero (i, j) con riga $i \in R$ ha colonna $j \in K$:
   - se non è in M, un cammino fino a i prosegue fino a j;
   - se è in M, la riga i è accoppiata, quindi è stata raggiunta dalla sua colonna j.

   Quindi L copre tutti gli zeri.

Dimostrazione (3/3)

5. *Conteggio.* Le righe non in R sono accoppiate, perché le righe non accoppiate stanno in R, e le colonne di K sono accoppiate.

   Uno zero di M in una colonna di K ha la sua riga in R, quindi sta esattamente su una linea di L.

   Linee e zeri di M si corrispondono uno a uno, e $|L| = |M|$.

□

- La dimostrazione è costruttiva: dà le linee di una copertura minima a partire da un insieme massimo di zeri indipendenti, e un insieme che non è massimo si allarga lungo un cammino alternante che termina in una colonna non accoppiata

Algorithm: L'algoritmo ungherese

Input : una matrice dei costi $n \times n$ c_{ij}, $i \in P$, $j \in D$
Output : un assegnamento ottimo e potenziali ottimi $\boldsymbol{\pi}$

1 **Passo 0 (riduzione).** Sottrai da ogni riga il suo minimo, π_i; poi sottrai da ogni colonna del risultato il suo minimo, $-\pi_j$. Sia $\bar{c}$ la matrice ridotta.
2 **Passo 1 (zeri indipendenti).** Trova un insieme massimo M di zeri indipendenti di $\bar{c}$.
3 Se $|M| = n$, `STOP`: M è un assegnamento ottimo.
4 **Passo 2 (linee).** Copri gli zeri con $|M| < n$ linee, come nella dimostrazione del teorema di König ed Egerváry.
5 **Passo 3 (aggiustamento).** Sia δ il minimo elemento non coperto. Sottrai δ dagli elementi non coperti e aggiungilo agli elementi coperti due volte: $\pi_i := \pi_i + \delta$ per ogni riga non coperta, $\pi_j := \pi_j + \delta$ per ogni colonna coperta. Torna al Passo 1.

Regola degli esempi e storia

- **Regola degli esempi.** Al Passo 1 l'insieme M si costruisce riga per riga, in ordine crescente: ogni riga cerca un cammino alternante verso una colonna non accoppiata, in profondità, provando le colonne in ordine crescente.

- L'algoritmo è dovuto a H. W. Kuhn (1955), che lo chiamò così in onore dei matematici ungheresi D. König e J. Egerváry.

Teorema 7: correttezza dell'algoritmo ungherese

A ogni iterazione $\bar{c}_{ij} = c_{ij} - \pi_i + \pi_j \geq 0$ *per ogni cella, e il Passo 3 aumenta il valore duale* $\sum_{i \in P} \pi_i - \sum_{j \in D} \pi_j$ *di* $\delta\,(n - k) > 0$, *dove* $k < n$ *è il numero di linee.*

Con costi interi l'algoritmo si ferma dopo un numero finito di iterazioni con un assegnamento ottimo, e $\boldsymbol{\pi}$ *è ottima per il duale del rilassamento LP.*

Dimostrazione (1/2)

1. *Passo 0.* Sottrarre il minimo di una riga lascia elementi non negativi e uno zero nella riga; i minimi di colonna sono allora ≥ 0, e sottrarli mantiene gli elementi non negativi. Per costruzione $\bar{c}_{ij} = c_{ij} - \pi_i + \pi_j$.

2. *Passo 3.* Tutti gli zeri sono coperti, quindi $\delta > 0$, e gli elementi non coperti restano ≥ 0. L'aggiornamento dei potenziali dà:
   - $-\delta$ su riga e colonna non coperte,
   - $-\delta + \delta = 0$ su riga non coperta e colonna coperta,
   - $+\delta$ su riga e colonna coperte,
   - 0 su riga coperta e colonna non coperta:

   esattamente l'aggiustamento.

Dimostrazione (2/2)

3. *Valore duale.* Con k_r righe coperte e k_c colonne coperte, il valore cresce di

$$\delta\,(n - k_r) - \delta\,k_c = \delta\,(n - k),$$

e $k = |M| < n$ per il teorema di König ed Egerváry.

4. *Terminazione.* Per il Lemma «costi ridotti di un assegnamento» il valore duale non supera mai il costo di un assegnamento. Con costi interi ogni δ è un intero ≥ 1, quindi il numero di aggiustamenti è finito. All'arresto M è un assegnamento di zeri, ottimo con π per il Corollario «gli assegnamenti a costo ridotto nullo sono ottimi».

□

Esempio: l'algoritmo ungherese (1/9)

Eseguiamo l'algoritmo sull'esempio dei quattro lavoratori.

Passo 0. I minimi di riga sono $2, 1, 2, 1$, quindi $\pi_1 = 2$, $\pi_2 = 1$, $\pi_3 = 2$, $\pi_4 = 1$, e le righe diventano

$c_{ij} - \pi_i$	5	6	7	8
1	0	2	1	3
2	3	7	0	0
3	0	4	7	4
4	0	8	7	5

Esempio: l'algoritmo ungherese (2/9)

- I minimi di colonna sono $0, 2, 0, 0$: cambia solo la colonna 6, e $\pi_6 = -2$
- Il valore duale $2 + 1 + 2 + 1 + 2 = 8$ è un limite inferiore al costo di ogni assegnamento

In ciascuna delle matrici che seguono:

- gli zeri indipendenti sono riquadrati,
- le linee sono blu,
- la cella di δ è rossa,
- i potenziali sono sui margini.

Esempio: l'algoritmo ungherese (3/9)

Iterazione 1

$\bar{c}_{ij}$	5	6	7	8	π_i
1	0	$\boxed{0}$	1	3	2
2	3	5	$\boxed{0}$	0	1
3	$\boxed{0}$	2	7	4	2
4	0	6	7	5	1
π_j	0	-2	0	0	

- Un insieme massimo di zeri indipendenti (riquadrati) ha 3 elementi: $(1,6)$, $(2,7)$, $(3,5)$.
- Le $3 < 4$ linee (blu) sono le righe 1 e 2 e la colonna 5.

Esempio: l'algoritmo ungherese (4/9)

Iterazione 1:

- Il minimo elemento non coperto è $\delta = 2$, nella cella $(3, 6)$ (rossa).
- Lo sottraiamo dagli elementi non coperti e lo aggiungiamo agli elementi coperti due volte: π_i cresce di 2 per le righe non coperte 3 e 4 e π_j cresce di 2 per la colonna coperta 5.
- Il valore duale cresce di $2 \cdot (4 - 3) = 2$.

Esempio: l'algoritmo ungherese (5/9)

Iterazione 2

$\bar{c}_{ij}$	5	6	7	8	π_i
1	2	$\boxed{0}$	1	3	2
2	5	5	$\boxed{0}$	0	1
3	$\boxed{0}$	0	5	2	4
4	0	4	5	3	3
π_j	2	-2	0	0	

- Un insieme massimo di zeri indipendenti (riquadrati) ha 3 elementi: $(1,6)$, $(2,7)$, $(3,5)$.
- Le $3 < 4$ linee (blu) sono la riga 2 e le colonne 5 e 6.

Esempio: l'algoritmo ungherese (6/9)

Iterazione 2:

- Il minimo elemento non coperto è $\delta = 1$, nella cella $(1, 7)$ (rossa).
- Lo sottraiamo dagli elementi non coperti e lo aggiungiamo agli elementi coperti due volte: π_i cresce di 1 per le righe non coperte 1, 3 e 4 e π_j cresce di 1 per le colonne coperte 5 e 6.
- Il valore duale cresce di $1 \cdot (4 - 3) = 1$.

Esempio: l'algoritmo ungherese (7/9)

Iterazione 3

$\bar{c}_{ij}$	5	6	7	8	π_i
1	2	0	[0]	2	3
2	6	6	0	[0]	1
3	0	[0]	4	1	5
4	[0]	4	4	2	4
π_j	3	-1	0	0	

- Un insieme massimo di zeri indipendenti (riquadrati) ha 4 elementi: $(1,7)$, $(2,8)$, $(3,6)$, $(4,5)$.
- Formano un assegnamento di costo ridotto nullo: l'algoritmo si ferma, e l'assegnamento è ottimo.

Esempio: l'algoritmo ungherese (8/9)

- L'assegnamento ottimo è

$$1 \to 7, \quad 2 \to 8, \quad 3 \to 6, \quad 4 \to 5,$$

con costo $3 + 1 + 6 + 1 = 11$, l'unico ottimo dell'esempio dei quattro lavoratori

- I potenziali finali $\pi_1 = 3$, $\pi_2 = 1$, $\pi_3 = 5$, $\pi_4 = 4$ e $\pi_5 = 3$, $\pi_6 = -1$, $\pi_7 = 0$, $\pi_8 = 0$ sono una soluzione duale ottima, di valore

$$(3 + 1 + 5 + 4) - (3 - 1 + 0 + 0) = 11.$$

- Il valore duale è cresciuto da 8 a 10 e poi a 11

Esempio: l'algoritmo ungherese (9/9)

Le linee sono quelle della dimostrazione del teorema di König ed Egerváry.

- Nell'iterazione 1 la riga 4 non è accoppiata
- Da essa un cammino alternante raggiunge la colonna 5 attraverso il suo zero, e poi la riga 3, accoppiata alla colonna 5
- Quindi $R = \{3, 4\}$, $K = \{5\}$, e le linee sono le righe 1, 2 e la colonna 5

Riepilogo

Il problema di assegnamento:

- è un problema di trasporto con offerte e domande unitarie
- il metodo dei potenziali vi fa molti pivot degeneri

L'algoritmo ungherese:

- mantiene $\bar{c}_{ij} = c_{ij} - \pi_i + \pi_j \geq 0$
- copre gli zeri con il numero minimo di linee (König–Egerváry)
- aumenta il valore duale finché n zeri indipendenti danno un assegnamento ottimo

Errori tipici

Introduzione: Errori tipici

- Raccogliamo gli errori più frequenti con i problemi di trasporto e di assegnamento
- Riguardano il bilanciamento dell'istanza, le basi, i potenziali, il pivot e l'algoritmo ungherese
- Controllali prima di svolgere gli esercizi

Errori tipici (1/4)

Osservazione 2: Errori tipici

- ***Risolvere un'istanza non bilanciata così com'è.*** *Con più offerta che domanda i vincoli di uguaglianza non hanno soluzione: aggiungere una destinazione fittizia a costi nulli, oppure un'origine fittizia quando la domanda supera l'offerta.*
- ***Togliere uno zero di base.*** *Una base degenere ha comunque $m + n - 1$ celle; senza la cella di valore 0 l'albero è sconnesso e i potenziali non si possono calcolare.*

Errori tipici (2/4)

Osservazione: Errori tipici

- ***Prendere come base un insieme con un ciclo.*** *$m + n - 1$ celle sono una base solo se non contengono un cammino chiuso che alterna righe e colonne.*
- ***Mescolare le convenzioni di segno.*** *Qui $\bar{c}_{ij} = c_{ij} - \pi_i + \pi_j$ con $\pi_i - \pi_j = c_{ij}$ sulle celle di base; con $u_i + v_j = c_{ij}$ i potenziali delle destinazioni cambiano segno.*

Errori tipici (3/4)

Osservazione: Errori tipici

- ***Calcolare θ su tutte le celle del ciclo.*** *θ è il minimo sulle sole celle con segno $-$; la cella entrante riceve θ.*
- ***Coprire gli zeri con troppe linee.*** *Le linee dell'algoritmo ungherese devono essere una copertura minima, tante quanti gli zeri indipendenti; con più linee δ può essere sbagliato o si può mancare l'arresto.*

Errori tipici (4/4)

Osservazione: Errori tipici

- ***Dimenticare gli elementi coperti due volte.*** *Al Passo 3 dell'algoritmo ungherese δ si sottrae dagli elementi non coperti e si aggiunge a quelli coperti da una riga e da una colonna.*
- ***Eseguire gli algoritmi sui profitti.*** *Gli algoritmi minimizzano: con profitti p_{ij} da massimizzare, usare i costi $c_{ij} = -p_{ij}$, oppure $c_{ij} = p_{\max} - p_{ij}$ con $p_{\max}$ il profitto più grande, per tenerli non negativi.*

Riepilogo

Prima di risolvere:

- bilanciare l'istanza con un'origine o una destinazione fittizia
- trasformare i profitti in costi

Metodo dei potenziali:

- tenere gli zeri di base, e non prendere mai come base un insieme con un ciclo
- una sola convenzione di segno; θ solo sulle celle con segno $-$

Algoritmo ungherese:

- una copertura minima, e δ aggiunto agli elementi coperti due volte

Riepilogo (1/3)

- Il modello TP spedisce le offerte a_i di m origini alle domande d_j di n destinazioni; è un modello MCF sulla rete bipartita completa, ammissibile esattamente quando è bilanciato
- Un'istanza di trasbordo senza capacità diventa un'istanza di trasporto con le distanze come costi
- Una base è un albero ricoprente della rete bipartita, con $m + n - 1$ celle; la sua soluzione di base si trova foglia per foglia ed è intera per dati interi

Riepilogo (2/3)

- Le regole dell'angolo nord-ovest e del costo minimo costruiscono una soluzione ammissibile di base, una linea per passo
- Il metodo dei potenziali: potenziali con $\bar{c}_{ij} = 0$ sulla base e $\pi_{m+n} = 0$; ottimo quando $\bar{c} \geq 0$; altrimenti θ unità si spostano lungo il ciclo della cella entrante, e il costo scende di $\theta\,|\bar{c}_e|$
- Il modello AP è un modello TP con dati unitari; l'algoritmo ungherese riduce righe e colonne, copre gli zeri con un numero minimo di linee (König–Egerváry) e aggiusta di δ, aumentando il valore duale

Riepilogo (3/3)

Le istanze in breve.

magazzini e negozi	nord-ovest 440, costo minimo 170, ottimo 165
metodo dei potenziali	4 pivot dal nord-ovest, 1 dal costo minimo
potenziali ottimi	$\tilde{\boldsymbol{\pi}} = (3, 2, 5, 2, 4, 1, 0)$, con $\pi_7 = 0$
con $a_1 = 40$	destinazione fittizia 8, ottimo 155
istanza degenere	$\theta = 0$, poi $\theta = 5$; ottimo 235
trasbordo senza capacità	ottimo 49
lavoratori e lavori	limite 8, $\delta = 2$, $\delta = 1$, ottimo 11