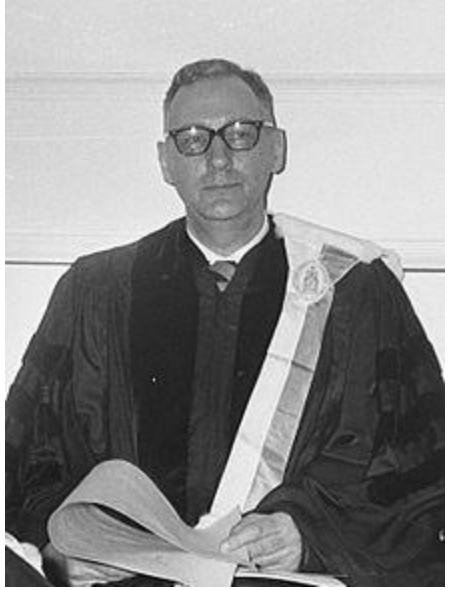
Transportation model: an introduction in GAMS

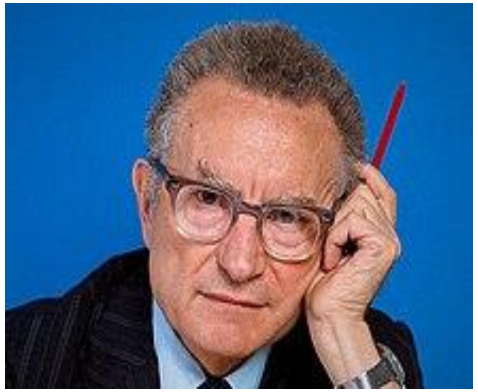
erwin@amsterdamoptimization.com

Transportation Model

- There is a deep connection with economic theory.
- Economists call this model: *Koopmans-Hitchcock Model*.
  - Tjalle Koopmans received the 1975 Nobel price in economics (with Kantorovich).
- Model 1 in the GAMS model library is a version of the transportation model in the 1963 George Dantzig book (originally a RAND report).
- Paul Samuelson noticed the connection between the transportation LP problem and the concept of spatial equilibrium.
  - Samuelson won the Nobel price in 1970.

Samuelson, Paul A. "Spatial Price Equilibrium and Linear Programming." *The American Economic Review* 42, no. 3 (1952)





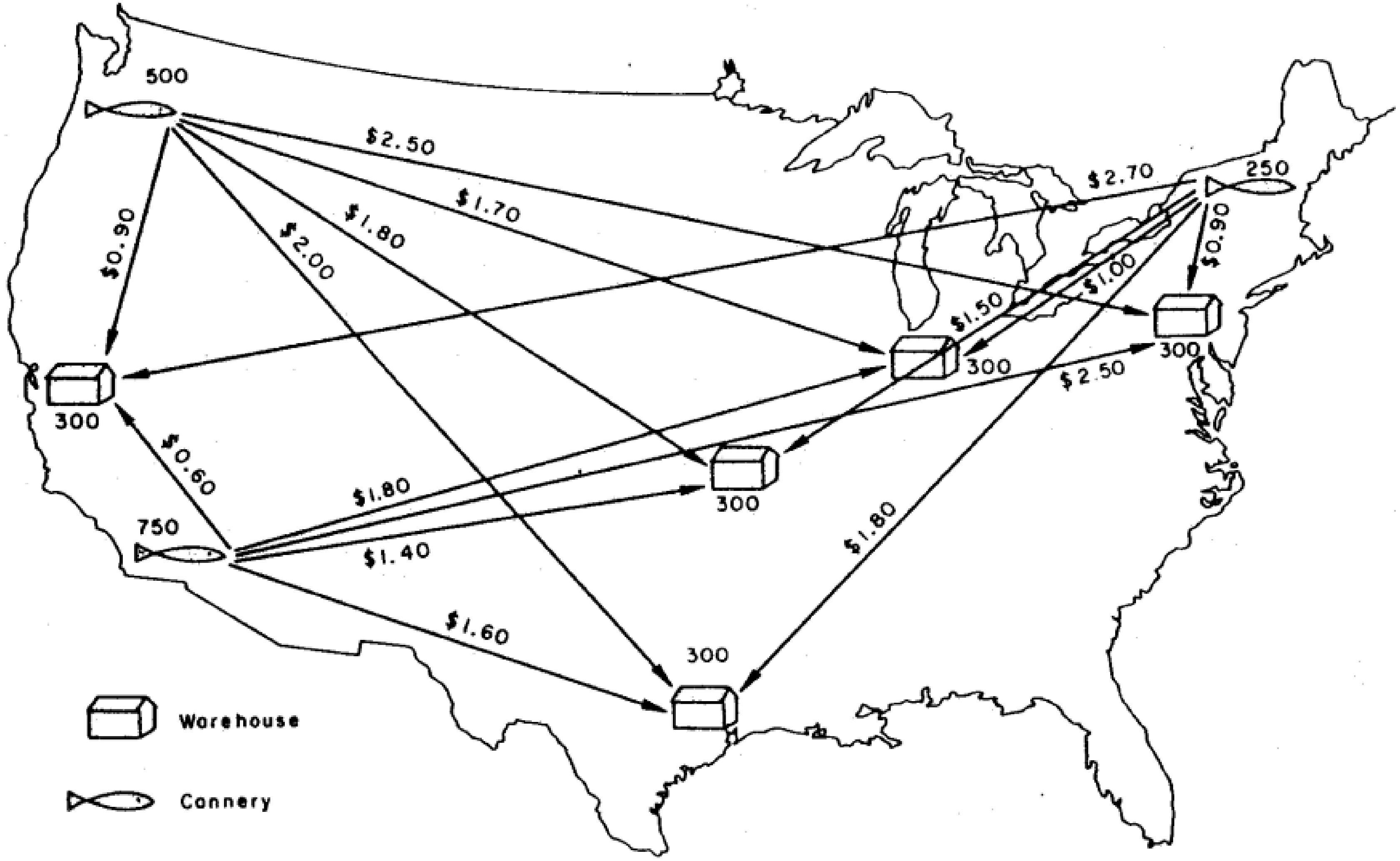
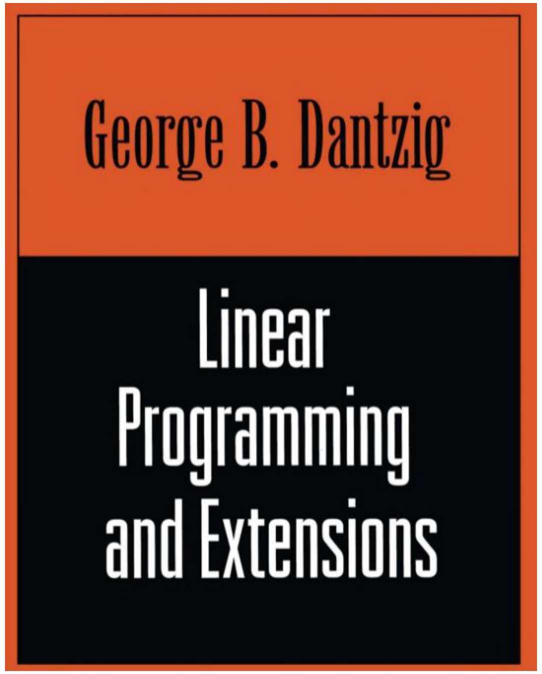
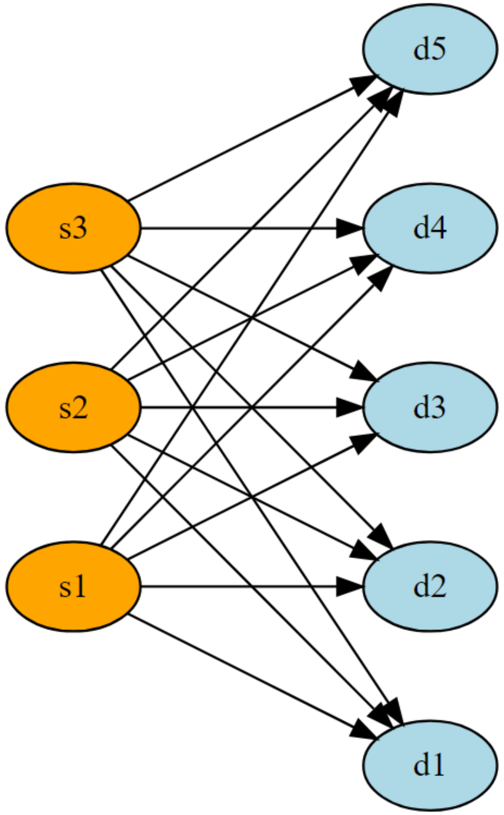


Figure 1-2-I. The Problem: Find a least cost plan of shipping from canneries to warehouses (the costs per case, availabilities and requirements are as indicated).

This picture is from the Dantzig book. The LP model is based on a smaller version discussed later in the book. The GAMS data is slightly different (maybe to make the solution degenerate).



Transportation model as network problem or LP.



Supply nodes

Demand nodes

$$
\begin{aligned}
\min \quad & \sum_{i,j} cost_{i,j} x_{i,j} \\
& \sum_{i} x_{i,j} \geq demand_j \quad \forall j \\
& \sum_{j} x_{i,j} \leq supply_i \quad \forall i \\
& x_{i,j} \geq 0
\end{aligned}
$$

Network problems have the property that each variable occurs exactly twice in the constraint matrix.

Model TRNSPORT from the model library, part 1

```
Set
   i 'canning plants' / seattle,  san-diego /
   j 'markets'        / new-york, chicago, topeka /;

Parameter
   a(i) 'capacity of plant i in cases'
        / seattle    350
          san-diego  600 /

   b(j) 'demand at market j in cases'
        / new-york   325
          chicago    300
          topeka     275 /;

Table d(i,j) 'distance in thousands of miles'
              new-york  chicago  topeka
   seattle         2.5      1.7     1.8
   san-diego       2.5      1.8     1.4;

Scalar f 'freight in dollars per case per thousand miles' / 90 /;

Parameter c(i,j) 'transport cost in thousands of dollars per case';
c(i,j) = f*d(i,j)/1000;
```

Set elements are strings (limit: 63 chars)

(1) Order not important
(2) Domain checked
(3) should use better names than a,b

Domain checking ensures referential integrity

Explanatory text is not ignored (like comments)

Excel, CSV files are popular external data sources

Parallel assignment

Don't use loops

Model TRNSPORT from the model library, part 2

```
Variable
   x(i,j) 'shipment quantities in cases'
   z      'total transportation costs in thousands of dollars';

Positive Variable x;

Equation
   cost      'define objective function'
   supply(i) 'observe supply limit at plant i'
   demand(j) 'satisfy demand at market j';

cost..      z =e= sum((i,j), c(i,j)*x(i,j));

supply(i).. sum(j, x(i,j)) =l= a(i);

demand(j).. sum(i, x(i,j)) =g= b(j);

Model transport / all /;

solve transport using lp minimizing z;

display x.l, x.m;
```

Default: variables are free

Blocks of variables

Blocks of equations

A model is collection of equations

GAMS uses an <u>objective variable</u> instead of an objection function
Calls external solver

.L: level, .M: marginal

Listing file has a lot of information

- Compilation output
  - Source listing
    - Useful to help with some syntax errors
- Execution time output
  - Output of DISPLAY statements
  - Output related to SOLVE statements
    - Model generation
      - Equation listing
        - Debug leads/lags
        - For non-linear models, Jacobian elements are shown
      - Column listing
        - Where does a variable appear?
      - Model statistics
        - Size of model
    - Solver messages
      - S O L V E   S U M M A R Y
        - Always check if solver succeeded
        - Model and Solver status
    - Solution listing
        - LO,L,UP,M

Number of variables and equations. Don’t forget that number of nonzero elements is also very important for sparse solvers!

Often users wonder about strange results without first checking the model and solver status

Exercises

1. Run the model and study the listing file.
2. Make a typo in one of the labels (set elements), and see how GAMS reacts.
3. Change the demand in NY to 400:
   ```
   b('new-york') = 400;
   ```
   (note that an element must be quoted). What happens?
4. Add check:
   ```
   abort$(totalDemand>totalSupply+0.0001) "Too much demand";
   ```
   If we pass this check, the model should be feasible.
5. Change demand in NY back to 325.
6. Pure network models have two non-zero elements in each column. Check the column listing for variable x. To view more columns in the column listing, add
   ```
   option limcol = 100;
   ```
   to the model. Notes:
   - the objective does not count
   - the default value for limcol is 3
   - a solver will typically substitute out the objective variable, and create an objective function.

LP,Marginals,Basis,EPS

- Marginals
  - **duals** for equations:
  - **reduced cost** for variables:
    - This is like a dual for $x_i \geq \ell_i, x_i \leq u_i$
  - Indicates: how much can obj change when bound/rhs changes
  - A marginal with value EPS means: numerically zero but this row/column is non-basic
    - Nerdy: Dual degeneracy (i.e. we can have multiple solutions)
    - See optional slides at end of this deck

Special values in GAMS: `EPS, INF, -INF, NA, UNDF`

Duals: how much can obj change

- Duals (marginal) of an equation indicates how much an objective can change when the rhs is increased.

```
---- EQU demand  satisfy demand at market j

              LOWER          LEVEL          UPPER         MARGINAL

new-york     325.0000       325.0000        +INF            0.2250
chicago      300.0000       300.0000        +INF            0.1530
topeka       275.0000       275.0000        +INF            0.1260
```

- Let's increase the demand in NY by 1. The obj can change by 0.225:

```
----    118 PARAMETER dualcheck  increase demand of NYC by one unit

            demand        dual         obj

before     325.000       0.225     153.675
after      326.000                 153.900
diff         1.000                   0.225
```

I.e. our marginal cost is 0.225. So we would want to charge NY at least 0.225 to meet this extra demand.

Duals: how much can obj change (cont'd)

```
parameter dualcheck(*,*) 'increase demand of NYC by one unit';

solve transport using lp minimizing z;

dualcheck('before','demand') = b('new-york');
dualcheck('before','dual') = demand.m('new-york');
dualcheck('before','obj') = z.l;

b('new-york') = b('new-york') + 1;
solve transport using lp minimizing z;

dualcheck('after','demand') = b('new-york');
dualcheck('after','obj') = z.l;
dualcheck('diff','demand') = b('new-york') - dualcheck('before','demand');
dualcheck('diff','obj') = z.l - dualcheck('before','obj');

display dualcheck;
```

Output is on previous slide

Exercises

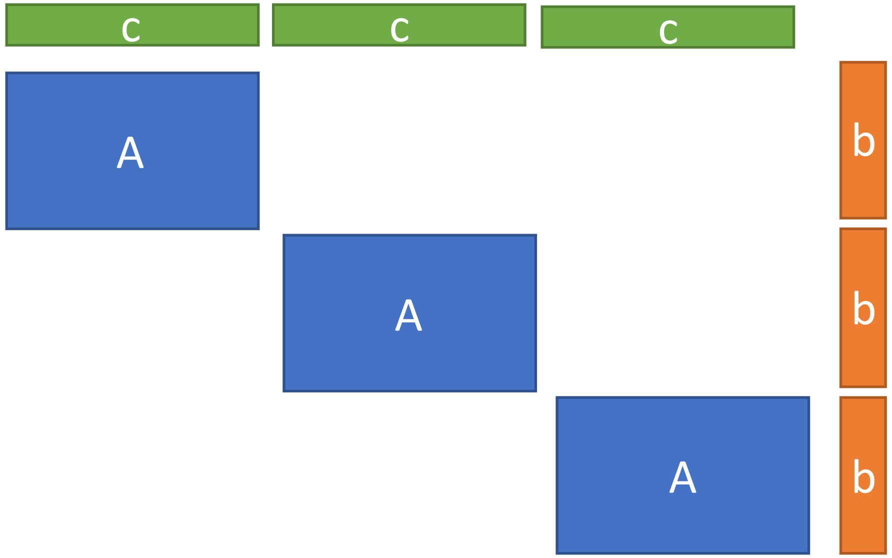
- What happens if we increase the demand in NY by 10 or 100.
- Add

  ```
  abort$(transport.modelstat <> %modelstat.optimal%) "Model was not solved to optimality";
  ```

  to alert about problems.
- Using the original NY demand of 325, would it make sense to increase capacity somewhere? (hint: look at the duals)

Exercise

- Make problem large by duplication
- Without inventing new data we can form:
- This can be viewed as a block-diagonal structure:



$$
\begin{aligned}
\min \quad & \sum_{i,j,k} cost_{i,j} x_{i,j,k} \\
& \sum_{i} x_{i,j,k} \ge demand_j \qquad \forall j,k \\
& \sum_{j} x_{i,j,k} \le supply_i \qquad \forall i,k \\
& x_{i,j,k} \ge 0
\end{aligned}
$$

```
set k /k1*k10/;
```

Try with 10, 100, 1000. These LPs should solve very fast

Adding an index to a symbol happens a lot in practical modeling (for many reasons)

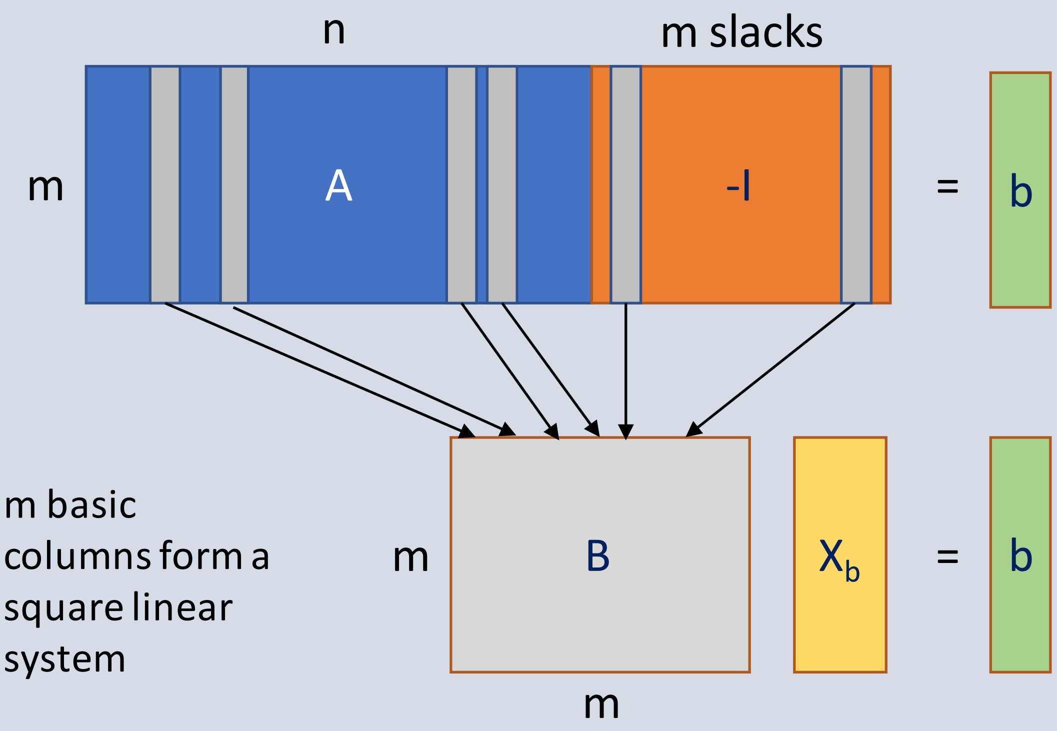
Optional

What is a basis?

Let m be the number of equations and n be the number of variables in out LP model.

In each Simplex iteration, n columns are temporarily fixed to one of their bounds. The remaining m columns form a square system which can be solved using linear algebra.

We say: there are m rows/columns that are **basic** and n rows/columns that are **non-basic**.



The m **basic** rows/columns have a value between their bounds and have a zero marginal (dual or reduced cost).

The n **non-basic** rows/columns are at one of the bounds and have a non-zero or EPS marginal.(*)

Exercise (use the listing file):

1. What is m,n in our transportation model?
2. Count the number of basic and non-basic rows/columns.
3. Check that non-basic rows/columns are at their bound.

* There is also something called superbasic. This is non-basic but between bounds. This is beyond our scope.

Optional

	LOWER	LEVEL	UPPER	MARGINAL	Basis status (concluded from marginal)
---- EQU cost	.	.	.	1.0000	NON-BASIC AT LOWERBOUND

cost define objective function

m=6 (equs)
n=7 (vars)

---- EQU supply observe supply limit at plant i

	LOWER	LEVEL	UPPER	MARGINAL	
seattle	-INF	350.0000	350.0000	EPS	NON-BASIC AT UPPERBOUND
san-diego	-INF	550.0000	600.0000	.	BASIC

---- EQU demand satisfy demand at market j

	LOWER	LEVEL	UPPER	MARGINAL	
new-york	325.0000	325.0000	+INF	0.2250	NON-BASIC AT LOWERBOUND
chicago	300.0000	300.0000	+INF	0.1530	NON-BASIC AT LOWERBOUND
topeka	275.0000	275.0000	+INF	0.1260	NON-BASIC AT LOWERBOUND

---- VAR x shipment quantities in cases

	LOWER	LEVEL	UPPER	MARGINAL	
seattle .new-york	.	50.0000	+INF	.	BASIC
seattle .chicago	.	300.0000	+INF	.	BASIC
seattle .topeka	.	.	+INF	0.0360	NON-BASIC AT LOWERBOUND
san-diego.new-york	.	275.0000	+INF	.	BASIC
san-diego.chicago	.	.	+INF	0.0090	NON-BASIC AT LOWERBOUND
san-diego.topeka	.	275.0000	+INF	.	BASIC

	LOWER	LEVEL	UPPER	MARGINAL	
---- VAR z	-INF	153.6750	+INF	.	BASIC

Basics: 6
Non-Basics: 7

Optional

Recognize an optimal LP solution

- The optimality conditions for an LP can be summarized as:
  1. The solution must be feasible
  2. The signs of the marginals must be:

Level	Minimization	Maximization
At Lower	≥ 0	≤ 0
At Upper	≤ 0	≥ 0

- If non-optimal GAMS will flag this:

```
---- VAR x  shipment quantities in cases
                          LOWER          LEVEL          UPPER          MARGINAL
seattle  .new-york          .              .            +INF            -0.0360  NOPT
seattle  .chicago           .           300.0000        +INF              .
seattle  .topeka            .            50.0000        +INF              .
san-diego.new-york          .           325.0000        +INF              .
san-diego.chicago           .              .            +INF             0.0450
san-diego.topeka            .           225.0000        +INF              .
```

Optional

EPS in solution

- A marginal=EPS means:
  - Numerically zero but non-basic
  - This means **dual degeneracy** (multiple optimal bases)
  - Note that different bases may only differ in marginals
    - All levels may stay the same
  - There is also something called **primal degeneracy**
    - This means: a basic variable is at its bound

Optional

Dual model

Primal

$$
\begin{aligned}
\min \quad & \sum_{i,j} cost_{i,j} x_{i,j} \\
& \sum_{j} x_{i,j} \le supply_i \perp u_i \le 0 \qquad \forall i \\
& \sum_{i} x_{i,j} \ge demand_j \perp v_j \ge 0 \qquad \forall j \\
& x_{i,j} \ge 0
\end{aligned}
$$

Dual

$$
\begin{aligned}
\max \quad & \sum_{i} supply_i \cdot u_i + \sum_{j} demand_j \cdot v_j \\
& u_i + v_j \le cost_{i,j} \perp x_{i,j} \ge 0 \qquad \forall i, j \\
& u_i \le 0, v_j \ge 0
\end{aligned}
$$

- $\perp$ means “with dual:”
- The signs on the duals are the optimality conditions
- Slight complication: `trnsport` has no unique optimal solution.

Exercise

1. Implement the dual model.
2. Compare results from primal.

Optional

Compare Results

Primal Model	Dual Model
VARIABLE **z**.L = 153.675	VARIABLE **zdual**.L = 153.675
---- 173 VARIABLE **x**.L *shipment quantities in cases* new-york chicago topeka seattle 50.000 300.000 san-diego 275.000 275.000	---- 173 EQUATION **e**.M *dual constraint* new-york chicago topeka seattle 50.000 300.000 san-diego 275.000 275.000
---- 173 EQUATION **supply**.M *observe supply limit at plant i* seattle EPS	---- 173 VARIABLE **u**.L *dual of supply* (ALL 0.000)
---- 173 EQUATION **demand**.M *satisfy demand at market j* new-york 0.225, chicago 0.153, topeka 0.126	---- 173 VARIABLE **v**.L *dual of demand* new-york 0.225, chicago 0.153, topeka 0.126