Videregående algoritmik

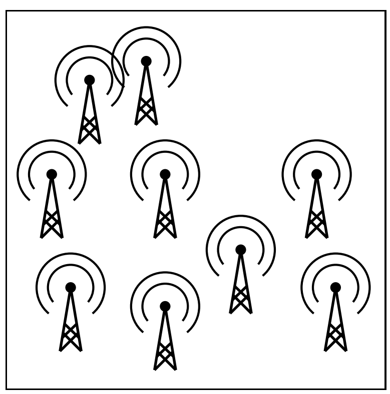
# Projektopgave 4

David Pisinger, Vinter 2004-05

Dette er den fjerde obligatoriske projektopgave på kurset “Videregående algorimik”. Opgaven stilles torsdag 6. januar 2005 og skal afleveres senest torsdag 13. januar 2005 kl. 12.00 i DIKU’s førstedelsadministration. For at blive godkendt skal der være gjort et reelt forsøg på at løse samtlige spørgsmål. Besvarelsen skal udarbejdes i grupper på to til tre deltagere. Grupper med en deltager kræver accept fra instruktoren. Læs venligst hele opgaveformuleringen igennem inden du går igang. Hints til opgaverne kan fås ved øvelserne, hvor der er afsat en øvelsesgang til at arbejde med projektopgaven.

## Kvadratisk knapsack problemet

Kvadratisk knapsack problemet har adskillige anvendelser. Betragt f.eks. et teleselskab som skal etablere et antal radiomaster rundt omkring i landet. De $n$ mest interessante byer udvælges, og der gennemføres en nøje analyse af hver radiomasts fordele og ulemper. En radiomast $j$ koster et givet beløb $w_j$ at etablere, og teleselskabet har et budget på $c$. Enhver radiomast vil i sig selv give et vist afkast $p_{jj}$ mens man får afkastet $p_{ij} + p_{ji}$ for telefoni mellem to byer $i$ og $j$.



Idet vi indfører beslutningsvariable $x_j \in \{0,1\}$ til at angive om en radiomast bygges eller ej, kan vi formulere problemet som følgende kvadratiske knapsack problem (QKP):

$$\begin{array}{ll} \text{maximize} & \sum_{j=1}^{n}\sum_{i=1}^{n} p_{ij}x_i x_j \\ \text{subject to} & \sum_{j=1}^{n} w_j x_j \le c \\ & x_j \in \{0,1\}, \quad j = 1,\ldots,n. \end{array} \tag{1}$$

hvor objektfunktionen angiver at profitten skal maximeres, mens knapsack begrænsningen betyder at etablerings-budgettet på $c$ ikke må overskrides. Det antages at alle værdier af $p_{ij}$ og $w_j$ er ikke-negative heltal.

Følgende eksempel viser en instans for syv byer.

| | $i \backslash j$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| $p_{ij}$ | 1 | 8 | 9 | 4 | 3 | 3 | 4 | 7 |
| | 2 | 9 | 7 | 1 | 2 | 6 | 3 | 9 |
| | 3 | 4 | 1 | 8 | 4 | 7 | 5 | 4 |
| | 4 | 3 | 2 | 4 | 1 | 0 | 3 | 7 |
| | 5 | 3 | 6 | 7 | 0 | 2 | 4 | 0 |
| | 6 | 4 | 3 | 5 | 3 | 4 | 8 | 9 |
| | 7 | 7 | 9 | 4 | 7 | 0 | 9 | 5 |
| $w_j$ | | 2 | 7 | 8 | 9 | 5 | 3 | 9 |

$n = 7, c = 20.$

Den optimale løsning er at bygge radiomaster i byerne $1, 2, 5, 6$ hvilket giver en profit på 83.

**Opgave 1** Formuler det tilhørende afgørlighedsproblem QKP-DECISION. ■

**Opgave 2** Vis at QKP-DECISION er $\mathcal{NP}$-fuldstændigt ved reduktion fra KNAPSACK. ■

## Øvre grænseværdi

**Opgave 3** Hvis man Lagrange-relaxerer knapsack begrænsningen i (1) fremkommer problemet $\mathrm{QKP}(\lambda)$ givet ved

$$\begin{array}{ll} \text{maximize} & \sum_{j=1}^{n} \sum_{i=1}^{n} p_{ij} x_i x_j - \lambda \left( \sum_{j=1}^{n} w_j x_j - c \right) \\ & x_j \in \{0, 1\}, \quad j = 1, \ldots, n. \end{array} \tag{2}$$

For hvilke værdier af $\lambda$ er (2) en relaxering af (1)? ■

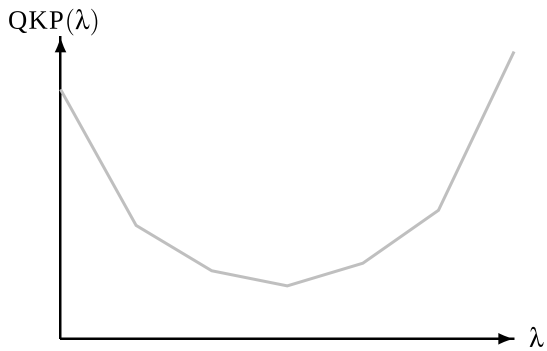
Betegn med $L$ den lovlige mængde af $\lambda$-værdier.

**Opgave 4** Vis at for et givet $\lambda \in L$ kan det Lagrange-relaxerede problem (2) skrives på formen

$$\begin{array}{ll} \text{maximize} & \left( \sum_{i \in N} \sum_{j \in N} d_{ij} x_i x_j \right) + k \\ \text{subject to} & x_j \in \{0, 1\}, \quad j = 1, \ldots, n \end{array} \tag{3}$$

hvor $k$ er en konstant. Vis endvidere at dette problem kan løses i polynomiel tid. ■

Det vides at $\mathrm{QKP}(\lambda)$ er en konveks funktion af $\lambda$, der ser ud som følger:




**Opgave 5** Det Lagrange duale problem går ud på at finde den værdi af $\lambda$ som resulterer i den strammeste grænseværdi. Dvs. vi søger

$$\min_{\lambda \in L} \text{QKP}(\lambda) \tag{4}$$

Konstruer og beskriv en algoritme som løser problemet (4). ■

**Opgave 6** Implementer algoritmen til løsning af (4) og afprøv den på instanserne fra hjemmesiden. ■

## Branch-and-bound algoritme

Vi vil nu konstruere en rekursiv branch-and-bound algoritme til løsning af QKP problemet (1). I hvert skridt forgrener vi på den sidste beslutningsvariabel $x_n$ således at det tilbageværende problem herefter har størrelse $n-1$.

Initielt sættes $x_i = 0$ for $i = 1, \ldots, n$ og $z^* = 0$. Herefter kaldes algoritmen

$$\texttt{quadknap}(0, c, n);$$

Den rekursive del af algoritmen kan skitseres som:

$\texttt{quadknap}(P, \overline{c}, \overline{n})$
  **if** $\overline{c} < 0$ **then return**
  **if** $P > z^*$ **then** $z^* \leftarrow P$; $x_i^* \leftarrow x_i$ for $i = 1, \ldots, n$
  **if** $\overline{n} = 0$ **then return**
  find en øvre grænseværdi $\mathcal{U}$ for problemet defineret på $(p_{ij})$, $i, j = 1, \ldots, \overline{n}$, og med kapacitet $\overline{c}$
  **if** $\mathcal{U} + P > z^*$ **then**
    $x_{\overline{n}} \leftarrow 1$; modificer $(p_{ij})$ matricen; $\texttt{quadknap}(P + p_{\overline{n}\overline{n}}, \overline{c} - w_{\overline{n}}, \overline{n} - 1)$
    $x_{\overline{n}} \leftarrow 0$; retabler $(p_{ij})$ matricen; $\texttt{quadknap}(P, \overline{c}, \overline{n} - 1)$

Her angiver $\overline{n}$ antallet af frie variable (dvs. variable som der endnu ikke er forgrenet på), $\overline{c}$ angiver den tilbageværende kapacitet, mens $P$ angiver profitten af de allerede valgte radiomaster.

Ved modification af $(p_{ij})$ matricen sætter vi

$$p_{ii} \leftarrow p_{ii} + p_{i\overline{n}} + p_{\overline{n}i}$$

for hvert $i = 1, \ldots, \overline{n} - 1$. Dette betyder, at hvis radiomast $i$ på et senere tidspunkt bliver valgt, vil man automatisk indkassere profitten af al kommunikation mellem byerne $i$ og $\overline{n}$. Når matricen skal retableres subtraherer vi atter disse bidrag ved at sætte

$$p_{ii} \leftarrow p_{ii} - p_{i\overline{n}} - p_{\overline{n}i}$$

for hvert $i = 1, \ldots, \overline{n} - 1$.

**Opgave 7** Implementer en fuldstændig version af `quadknap` som benytter grænseværdien (4). ■

**Opgave 8** Kør algoritmen på instanserne fra hjemmesiden og angiv køretid, antal branch-and-bound knuder, grænseværdi i rodknuden, samt optimal løsningsværdi. ■

## Noter

Til løsning af de sidste opgaver benyttes rammeprogrammet fra projektopgave 3. Rammeprogrammet findes på kursets hjemmeside sammen med alle instanser.

De benyttede datatilfælde er en blanding af virkelige data og tilfældigt genererede data som angivet i følgende tabel:

| Beskrivelse | filnavne |
|---|---|
| Radiotelefoni | `tele7` |
| Compiler design | `comp30 comp45 comp47` |
| Geometrisk $p$-dispersion | `geo10 geo20 geo30 geo40` |
| Random 25% densitet | `rand10.25 rand20.25 rand30.25 rand40.25` |
| Random 100% densitet | `rand10.100 rand20.100 rand30.100 rand40.100` |
| Klike i en graf | `clique10 clique20 clique30 clique40` |

Densiteten af et datatilfælde angiver hvor mange procent af $p_{ij}$'erne der er forskellige fra nul.