Vejledning

Implementeringsopgaver er markerede med *. Vi anbefaler at I begynder med at besvare de teoretiske spørgsmål som giver grundlag for algoritmerne.

Sparse løsninger til lineære ligningsystemer

I flere problemer i machine learning (ML) og compressed sensing er man interesseret i at finde løsninger til et system af algebraiske ligninger, som kun har få indgange forskellig fra 0. Dette kan gøres ved at løse et optimeringsproblem:

$$\begin{aligned} \text{minimize} \quad & \sum_{i=1}^{n} |x_i|, \\ \text{subject to} \quad & Ax = b \end{aligned} \tag{1}$$

hvor matricen A, med m rækker og n søjler, og vektoren $b \in \mathbb{R}^m$ er givet.

1. Problemet i (1) er ikke et lineært optimeringsproblem da objektfunktionen $\sum_{i=1}^{n} |x_i|$ ikke er en lineær transformation fra $\mathbb{R}^n$ til $\mathbb{R}$. Forklar med (mod)eksempel, helst i dimension $n = 1$, hvorfor objektfunktionen ikke opfylder definition af en lineær transformation på side 82 i [Lay].

2. Lad os betragte en situation hvor $n = m = 1$, $A = [2]$, $b = [\beta]$. Bestem en løsning til (1) som en funktion af $\beta \in \mathbb{R}$. Tegn mulighedsområdet $\mathcal{F}$ og bestem en løsning til følgende lineære optimeringsproblem

$$\begin{aligned} \text{minimize} \quad & x^+ + x^-, \\ \text{subject to} \quad & \\ & 2x^+ - 2x^- = \beta, \\ & x^+ \geq 0, x^- \geq 0, \end{aligned} \tag{2}$$

for $\beta \geq 0$ og $\beta < 0$. Konkluder at hvis $(\bar{x}^+, \bar{x}^-)$ er en løsning til (2), så er $\bar{x} = \bar{x}^+ - \bar{x}^-$ en løsning til (1).

For større dimensioner n, m, er idéen i det sidste spørgsmål stadig gyldig. Dvs, lad

$$\tilde{x} = \begin{bmatrix} x^+ \\ x^- \end{bmatrix} \in \mathbb{R}^{2n}, \qquad \tilde{c} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \in \mathbb{R}^{2n}, \qquad \tilde{A} = \begin{bmatrix} A & -A \end{bmatrix},$$

hvor vektorerne $x^+ \in \mathbb{R}^n$, $x^- \in \mathbb{R}^n$. Da er (1) ækvivalent med et lineært optimeringsproblem:

$$\begin{aligned} \text{minimize} \quad & \tilde{c} \cdot \tilde{x}, \\ \text{subject to} \quad & \\ & \tilde{A}\tilde{x} = b, \\ & \tilde{x} \geq 0. \end{aligned} \tag{3}$$

3. Omskriv problemet (3) til et lineært problem i kanonisk form, og bestem det duale problem.

4. Antag at $m = 1$, $n = 5$, $A = [1, 2, 3, 4, 5]$, $b = [10]$. Ved hjælp af en tegning find en løsning til det duale problem I har bestemt i det sidste spørsmål. Brug stærk dualitet til at bestemme en løsning til det primale problem ud fra løsningen til det duale. Løs også det primale problem vha simplexmetoden[1], og tjek at I får samme resultat.

5. Lad os betragte mulighedsområdet til (3) uden at omskrive problemet til kanonisk form. Vi arbejder stadig med $A = [1, 2, 3, 4, 5]$, $b = [10]$. Find alle 5 mulige basisløsninger til systemet $\tilde{A}\tilde{x} = b$, $\tilde{x} \geq 0$. (Disse løsninger svarer til "hjørne"/ekstremapunkter af det 10-dimensionelle mulighedsområde for $\tilde{x}$, som er selvfølgelig ikke så nemt at visualisere.) Find objektfunktionens værdi i disse 5 punkter og sammenlign med jeres svar fra forrige opgave.

6. * Funktion `solveLP(c,Ain,bin,Aeq,beq)` fra pakken `numeric` kan bruges til at løse lineære optimeringsproblemer på formen:

$$\begin{aligned} \text{minimize} \quad & c \cdot x, \\ \text{subject to} \quad & \\ & A_{\text{in}} x \leq b_{\text{in}}, \\ & A_{\text{eq}} x = b_{\text{eq}}. \end{aligned}$$

[1] Gennemgå exempel 7 i afsnit 9.3 for å gøre det

Advarsel: alle uligheds-bibetingelsene, inklusive de mulige fortegnsrestriktioner, skal omskrives på formen $A_{\text{in}}x \leq b_{\text{in}}$. Brug denne funktion til at løse problemet fra delspørsmål 4–5 numerisk, og sammenling resultaterne med dem, I havde beregnet:

```
// forbered Ain,bin,Aeq,beq,c, som arrays
// loes problemet
var lp = numeric.solveLP(c,Ain,bin,Aeq,beq);
// afrund loesningen
var solution = numeric.trunc(lp.solution,1.0E-12);
```

Tjek at array `solution` indenholder en numerisk tilnærmelse til en løsning.

Del 2. Anvendelse: sparse sensing

Antag at vi måler et tidsafhængigt signal $f(t)$ i diskrete tidspunkter $t_1, t_2, \ldots, t_m$. Vi vil gerne finde en repræsentation af f som en lineær kombination af visse "basisfunktioner" $g_1, g_2, \ldots, g_n$, $n \geq m$, således at:

$$f(t_i) = \sum_{j=1}^{n} x_j g_j(t_i), \qquad \forall i = 1, \ldots, m,$$

og at $\sum_{j=1}^{n} |x_j|$ er mindst muligt.

1. Formuler problemet så det har samme form som i (1); dvs, bestem en matrix A ved hjælp af g_j og t_i, og den højre side b ved hjælp af f og t_i.

2. * Antag nu at $m = 5$, $n = 20$, $t_i = (i-1)/m$, $i = 1, \ldots, m$, $f(t) = 1/(25+t^2)$, $g_j(t) = \cos((j-1)\arccos(t))$, $j = 1, \ldots, n$. Løs problemet numerisk ved hjælp af `solveLP` som i den tidligere del, og bekræft, at mange indgange i løsningen $\bar{x}$ til (1) er lig 0. Sammenlign $f(t)$ og $\sum_{j=1}^{n} \bar{x}_j g_j(t)$ ved at plotte begge funktioner på intervallet $(0, 1)$. (Dette kan gøres f.eks. i Maple eller Matlab. Hvis I har en god ide, hvordan dette nemt kan gøres i Javascript, vil jeg gerne høre om det.)