Chalmers/GU
Matematiska institutionen

TENTAMEN

TMA945
TILLÄMPAD OPTIMERINGSLÄRA

Datum:	98–08–19
Tid:	8^{45}–13^{45}
Hjälpmedel:	Typgodkänd räknedosa
Antal uppgifter:	7; för godkänd uppgift krävs 2 poäng av 3. Uppgifterna är *inte* ordnade efter svårighetsgrad. För godkänt krävs 8 poäng och tre godkända uppgifter
Examinator:	Michael Patriksson
Jourhavande lärare:	Sebastian Sandberg, tel. 5352
Resultat anslås senast:	98–08–31 Kortfattade lösningsförslag anslås vid skrivningstidens slut på anslagstavlan för optimeringslära vid matematiska institutionen.

Tentamensinstruktioner

När Du löser uppgifterna

Motivera och redovisa noga lösningsmetodik och beräkningar.
Använd de standardmetoder som genomgåtts på föreläsningar och lektioner.

Skriv endast på ena sidan av lösningsbladen. Använd inte rödpenna.
Behandla ej fler än en huvuduppgift på varje blad.

Vid skrivningens slut

Sortera Dina lösningsblad i uppgiftsordning.
Markera på omslaget de uppgifter Du behandlat.
Kontrollräkna antalet inlämnade blad och fyll i antalet på omslaget.

(3p) Uppgift 1

En ny producent av parfym önskar slå sig in på en lukrativ marknad. En exklusiv parfym, Chinelle, skall tillverkas och marknadsföras till maximal intäkt. Med den tillgång till maskiner man har kan man producera varan i två olika slags processer, och man funderar också på att anlita en fotomodell för att lansera den. För att förenkla problemställningen, låt oss anta att parfymen tillverkas med hjälp av två huvudingredienser, ämnet med den hemliga formeln MO och en blandning av allmänt förekommande ämnen. Två processer är tänkbara, med vilka man kan tillverka produkten. Den första ger tre gram parfym av en enhet MO och två enheter av det andra ämnet, medan den andra processen ger fem gram parfym av två respektive tre enheter av de två ämnena. Man disponerar över tillverkningsprocesser som kan tillverka 20,000 enheter MO under planeringshorisonten och 35,000 enheter av den andra blandningen. Varje enhet MO kostar tre FFr (franska francs, vi använder naturligtvis den centrala valutan i sammanhanget) att tillverka, och den andra blandningen två FFr per enhet. Ett gram parfym beräknas kunna säljas för femtio FFr. Utan marknadsföring tror man sig kunna sälja 1,000 gram parfym enbart genom nyhetens behag, men man vill höja efterfrågan med hjälp av en annonskampanj. Därtill finns en fotomodell som ställer upp för fotosessioner för priset 5,000 FFr per halvtimme, och man tror att en kampanj kan höja efterfrågan med motsvarande 200 gram per halvtimme utnyttjad tid, upp till maximalt tre timmar. Formulera problemet att välja den optimala strategin som ett LP-problem.

(3p) Uppgift 2

Stora M-metoden med en (1) artificiell variabel är en enkel teknik för att snabbt hitta en tillåten baslösning för simplex-metoden, och som undviker lösandet av ett Fas I-problem. För problemet (här är A en $m \times n$-matris)

$$[\text{P}] \qquad \begin{aligned} \min \quad & f(x) := c^{\mathrm{T}}x, \\ \text{då} \quad & Ax = b, \\ & x \geq 0^n, \end{aligned}$$

går det till så att man först finner en godtycklig (och troligen otillåten) baslösning (x_B, x_N), varefter man adderar en och samma artificiella variabel $a \geq 0$ med koefficient -1 på de rader för vilka motsvarande element i $x_B := B^{-1}b$ är negativa. Variabeln a ges en målfunktionskoefficient, M, vars värde är stort nog för att den inte skall kunna ingå i en optimal baslösning till det nya problemet. Pivoteras a in i basen så att den ersätter den mest negativa basvariabeln i x_B blir den nya baslösningen tillåten. Därefter kan vi alltså lösa problemet med simplexmetoden som vanligt.

Vi är nu intresserade av att använda en liknande teknik för den *duala simplex-metoden* (som vi minns som ett sätt att lösa det duala problemet via primala simplexpivoteringar). Er uppgift är att härleda en lika enkel startmetod för denna metod som den vi just har presenterat för den primala simplexmetoden. dvs. beskriva en teknik som enkelt finner en *dualt* tillåten baslösning.

Tips: om man applicerar stora M-metoden på det duala problemet till [P], vilken tolkning kan detta ges i termer av [P] självt?

(3p) **Uppgift 3**

Betrakta optimeringsproblemet

[P]
$$\begin{aligned} \min \quad & f(x), \\ \text{då} \quad & x \in \Re^n, \end{aligned}$$

där $f : \Re^n \mapsto \Re$ är kontinuerligt differentierbar.

Låt $\{x^t\}$ vara en sekvens av iterationspunkter som genererats av någon algoritm för lösandet av [P], och antag att för denna sekvens gäller att $\{\nabla f(x^t)\} \to 0^n$, alltså att gradienten konvergerar mot noll (vilket är ett önskvärt beteende). Den aktuella frågan är vad det betyder för den eventuella konvergensen hos $\{x^t\}$.

Betrakta alltså sekvensen $\{x^t\}$ och också sekvensen $\{f(x^t)\}$ av målfunktionsvärden. Givet antagandet att $\{\nabla f(x^t)\} \to 0^n$, gäller det att $\{x^t\}$ eller $\{f(x^t)\}$ konvergerar eller ens är begränsade? Nämn samtliga fall som kan förekomma när det gäller konvergensen hos dessa sekvenser. Ange ett exempel, gärna för $n = 1$, för varje fall.

Uppgift 4

Betrakta LP-problemet (A är en $m \times n$ matris)

[P]
$$\begin{aligned} \min \quad & f(x) := c^{\mathrm{T}} x, \\ \text{då} \quad & Ax = b, \\ & x \geq 0^n. \end{aligned}$$

Betrakta en tillåten baslösning (x_B, x_N) till [P]. Från denna punkt i $\Re^n$ utgår $n - m$ halvlinjer motsvarande den förflyttning som sker då de $n - m$ icke-basvariablerna i den aktuella basen ökar sina respektive värden från noll. [Om baslösningen (x_B, x_N) är icke-degenererad beskriver dessa halvlinjer precis de $n - m$ kantlinjerna till polyedern som utgår från extrempunkten som motsvaras av (x_B, x_N), och ger alltså en koppling mellan kantlinje (geometri) och icke-basvariabel (algebra); i en degenererad baslösning är inte alla halvlinjer kantlinjer till polyedern, några leder omedelbart ut ur polyedern.] Dessa halvlinjer har riktningarna $p_i \in \Re^n$ $(i = 1, \ldots, n - m)$.

(1p) a) Visa att det traditionella kriteriet vid valet av den inkommande basvariabeln motsvarar precis att välja den halvlinje som lutar brantast, dvs. att välja det i i $\{1, 2, \ldots, n - m\}$ som ger lägst värde $c^{\mathrm{T}} p_i$.

(2p) b) Med ledning av denna geometriska tolkning, ange minst en nackdel med det traditionella kriteriet och föreslå ett alternativt kriterium för valet av den inkommande basvariabeln som eliminerar denna nackdel.

(3p) Uppgift 5

I fallet obegränsad, konvex optimering vet vi att optimallösningarna kan ges en elegant karakterisering som de vektorer x^* för vilka gradientens värde är lika med noll. Denna uppgift syftar till att visa att en motsvarande slutsats kan dras också för *begränsad*, konvex optimering. I detta mer generella fall visar vi att gradientens värde hos varje optimallösning är lika (men ej troligen lika med noll som fallet är i obegränsad optimering). [Detta innebär bland annat att algoritmer för konvex, begränsad optimering kan utnyttja förändringarna hos gradientens värde mellan successiva iterationer i avbrottskriterier.]

Betrakta optimeringsproblemet

$$[\mathrm{P}] \qquad \begin{aligned} \min \quad & f(x), \\ \text{då} \quad & x \in X, \end{aligned}$$

där $f : \Re^n \mapsto \Re$ är konvex och kontinuerligt differentierbar, och där $X \subseteq \Re^n$ är en icke-tom konvex mängd. Antag att X^* är mängden av globala optimallösningar till [P]. Vi antar att denna mängd är icke-tom (dvs. att problemet [P] har en optimal lösning) och att ett element i X^* är vektorn x^*.

Låt nu

$$\hat{X} := \{\, x \in X \mid \nabla f(x^*)^{\mathrm{T}}(x - x^*) = 0; \quad \nabla f(x) = \nabla f(x^*) \,\}.$$

Visa att $X^* = \hat{X}$.

Tips: utnyttja vad Ni vet om primala optimalitetsvillkor (och förstås ekvivalenta definitioner av konvexitet hos en funktion).

Uppgift 6

Ibland är beslutsfattare omedvetna om begränsningarna hos optimeringsmodeller. Ett exempel på något som en beslutsfattare ofta vill göra är att optimera flera mål samtidigt. Det är därför inte förvånande att ett flertal tekniker inom flermålsoptimering har utvecklats, och vi skall titta på en interaktiv sådan inom linjär optimering.

Antag att en beslutsfattare är intresserad av att lösa ett LP-problem med två olika målfunktioner. Vi vet dock att han/hon har en inneboende preferens (som kanske inte kan uttryckas i klartext), vilken vi simulerar genom att generera flera lösningalternativ. Dessa skapas som "kompromisslösningar", där de två funktionerna vägs samman i en konvexkombination, som varieras över alla möjliga värden. Optimallösningarna presenteras sedan för beslutsfattaren, som sedan kan välja bland dem enligt sin preferens.

Betrakta alltså polyedern $X := \{\, x \in \Re^n \mid Ax = b; \quad x \geq 0^n \,\}$, och två målfunktioner, $f^1(x) := (c^1)^{\mathrm{T}}x$ och $f^2(x) := (c^2)^{\mathrm{T}}x$, som skall maximeras över X. Kompromisslösningar fås genom att maximera funktionen $\lambda f^1(x) + (1-\lambda)f^2(x)$, dvs. $[\lambda c^1 + (1-\lambda)c^2]^{\mathrm{T}}x$, över $x \in X$, där λ varieras kontinuerligt inom intervallet $[0,1]$.

(2p) a) Beskriv hur samtliga kompromisslösningar kan genereras med hjälp av simplexmetoden.

Tips: utnyttja Era kunskaper i känslighetsanalys.

(1p) b) Illustrera tekniken på följande data:

$$A = \begin{pmatrix} 1 & 1 & 1 & 0 \\ -1 & 2 & 0 & 1 \end{pmatrix}; \qquad b = \begin{pmatrix} 6 \\ 6 \end{pmatrix}; \qquad c^1 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}; \qquad c^2 = \begin{pmatrix} -3 \\ 1 \end{pmatrix}.$$

Notera: här är det tillåtet att använda geometrisk lösning för den som så önskar.

Uppgift 7

Betrakta optimeringsproblemet

[P]
$$\begin{aligned} \min \quad & f(x), \\ \text{då} \quad & x \in \Re^n, \end{aligned}$$

där $f : \Re^n \mapsto \Re$ är två gånger kontinuerligt differentierbar.

(1p) a) Låt $x^t \in \Re^n$ och betrakta följande två riktningsbestämmande problem (Δ och γ är positiva konstanter och $\|\cdot\|$ betecknar Euklidisk norm):

$$\begin{aligned} \min \quad & \nabla f(x^t)^{\mathrm{T}} p, \\ \text{då} \quad & \|p\| \leq \Delta; \end{aligned}$$

$$\begin{aligned} \min \quad & \nabla f(x^t)^{\mathrm{T}} p + 1/(2\gamma)\|p\|^2, \\ \text{då} \quad & p \in \Re^n. \end{aligned}$$

Visa att dessa är ekvivalenta i den meningen att optimallösningarna p har samma riktning, oavsett valet av $\Delta > 0$ och $\gamma > 0$. Vilken metod bygger på dessa riktningsbestämmande problem?

(2p) b) Ett alternativ till "sökmetoder" (där man alternerar mellan lösandet av riktningsbestämmande problem, á lá det andra alternativet ovan, och lösandet av endimensionella minimeringsproblem) är en "trust region"-ansats, á lá det andra alternativet ovan. Där använder man "modeller" av f (dock inte linjära som ovan) som minimeras över en (enkel) begränsad mängd, oftast en boll, där modellen anses som "tillförlitlig". Beroende på hur stor förbättring i målfunktionsvärdet som lösningen ger jämfört med en given iterationspunkt används lösningen för att definiera en ny iterationspunkt, och bollen görs större eller mindre beroende på modellens tillförlitlighet. Ett exempel på en sådan metod, baserad på en modell á lá Newtons metod, ges nedan.

0. (*Initialisering*): Välj en startpunkt $x^0 \in \Re^n$. Välj $\Delta_0 > 0$, $0 < \mu < \nu < 1$. Sätt $t := 0$.

1. (*Minimering av modellen*): Lös problemet

$$\begin{array}{rl} \min & \psi_t(p) := f(x^t) + \nabla f(x^t)^{\mathrm{T}} p + (1/2) p^{\mathrm{T}} \nabla^2 f(x^t) p, \\ \text{då} & \|p\| \le \Delta_t. \end{array}$$

   Låt p^t beteckna en optimallösning till detta problem.
2. (*Test av modellens tillförlitlighet*): Beräkna

$$r_t := \frac{f(x^t) - f(x^t + p^t)}{f(x^t) - \psi_t(p^t)},$$

   dvs. kvoten mellan den verkliga och modellens predikterade förbättring av målfunktionsvärdet. Låt

$$x^{t+1} := \begin{cases} x^t, & \text{om } r_t \le \mu \text{ (misslyckad iteration)}, \\ x^t + p^t, & \text{om } r_t > \mu \text{ (lyckad iteration)} . \end{cases}$$

3. (*Uppdatering av bollens storlek*): Låt

$$\Delta_{t+1} := \begin{cases} \frac{1}{2}\Delta_t, & \text{om } r_t \le \mu \text{ (liten tillförlitlighet)}, \\ \Delta_t, & \text{om } \mu < r_t < \nu \text{ (hyfsad tillförlitlighet)}, \\ 2\Delta_t, & \text{om } r_t \ge \nu \text{ (stor tillförlitlighet)}. \end{cases}$$

   Låt $t := t + 1$ och gå till Steg 1.

Visa att p^t löser det linjära ekvationssystemet (I^n är enhetsmatrisen i $\Re^{n \times n}$)

$$[\nabla^2 f(x^t) + \lambda I^n] p = -\nabla f(x^t)$$

för något $\lambda \ge 0$ där $\nabla^2 f(x^t) + \lambda I^n$ är positivt semidefinit.

Antag nu att $\nabla^2 f(x^t)$ är positivt definit, och utred hur p varierar då λ varierar mellan 0 och $+\infty$. Med ledning av detta, tolka metoden ovan i termer av sökmetoder och ange en motivering till när den är att föredra framför sökmetoder.