



EXAMENSARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

Dualitetsmetoder i stokastisk analys

av

Karin Grankvist och Janne Åkerström

2008 - No 15

Dualitetsmetoder i stokastisk analys

Karin Grankvist och Janne Åkerström

Examensarbete i matematik 30 högskolepoäng, fördjupningskurs

Handledare: Yishao Zhou

2008

Sammanfattning

Denna uppsats behandlar studiet av optimeringsproblem, särskilt stokastiska sådana. Vi studerar de fascinerande och kraftfulla resultat som dualitetsmetoder för optimering applicerad på stokastiska miljöer ger. Vi inleder med att grundligt förklara och sedan fördjupa oss i de i sig åtskilda delarna dualitetsteori och Itökalkyl, för att i slutet visa att det är först när de används tillsammans som dessa metoder för optimering kommer till sin fulla rätt.

⁰Kontaktinformation: janne@akerstrom.nu & karingrankvist@gmail.com

Innehåll

1	Inledning	5
2	Bakgrund och kort diskussion	7
3	Den Brownska Rörelsen	9
4	It\bar{o}s integral	12
4.1	Stokastiska integraler över funktioner	14
5	It\bar{o}s formel och -process	16
5.1	It \bar{o} processer	16
5.2	Förenklad notation för It \bar{o} processer	16
5.3	Martingaler	17
5.4	It \bar{o} s formel	18
5.5	It \bar{o} s formel i n-dimensionella fallet	19
5.6	Några beräkningstekniska resultat	20
6	Stokastiska differentialekvationer	24
6.1	Uppställning av stokastisk differential ekvation, SDE	24
6.2	Lösning av enkla SDEs	25
6.3	Lösning av "svårare" SDEs	25
7	Inledning till dualitet	29
8	Lagrangedualitet	30
8.1	Svag dualitet	32
8.2	Dualitetsgap	32
8.3	Stark dualitet	36
9	Sadelpunktskriterier och KKT-villkoren	39
9.1	Sadelpunktskriterier	39
9.2	Karush-Kuhn-Tuckers villkor	40
9.3	Relationen mellan sadelpunktsoptimalitet och KKT-villkoren	42
10	Duala funktionens egenskaper	45
11	Primala och duala problemen	53
11.1	Att formulera det duala problemet	53
11.2	Skärningsplan- eller yttre linjäriseringsmetoden	53
11.3	Primala problemet	57

12 Konjugatdualitet	58
12.1 Konjugerade funktioner	59
12.2 Fenchels dualitetssatser	61
13 Linjär-kvadratisk kontroll och dualitet	63
14 Stokastisk kontroll	66
15 Tillämpningar av icke-linjära problem	69
15.1 Stokastisk tillgångsallokering	69
15.2 Stokastisk programmering	71
16 Finansiella applikationer	72
16.1 Marknaden	72
16.2 Fordringar	75
17 Dualitet och portföljoptimering	82
18 Datorsimulering	87
18.1 Brownsk rörelse	87
18.2 m-dimensionell Brownsk rörelse	87
18.3 Den stokastiska integralen	88
18.4 Itoprocesser	88
18.5 Black-Scholes prisformel för optioner	89
A Appendix1: Tabeller med Matlab-simuleringars värden	92
A.1 Tabell över stokastisk integral	92
A.2 Tabell över skattad Itoprocess	92
A.3 Tabeller över skattade optionspriser	92
B Appendix 2: Matlab-koder	94
C Appendix3: Bilder	98

1 Inledning

Matematiska modeller används dagligen för att förklara olika fenomen som vi människor stöter på i världen. Vi använder matematiken för att räkna ut hur vi skall gå till väga för att få optimala resultat av våra handlingar och vi använder den för att i efterhand få förståelse för hur ett skeendes förlopp såg ut. I många fall, särskilt när vi söker ett a priori-råd för hur vi skall gå till väga, uppkommer problemet att världen alltid har ett element av slumpmässighet inom sig. Vi som studenter på matematik-ekonomilinjen är naturligtvis främst intresserade av ekonomiska fenomen, något som kan tyckas skulle hjälpa oss att komma ifrån problemet med osäkra utfall. Det nationalekonomiska antagandet att alla människor är rationella borde undanröja alla inslag av slump i de beräkningar vi önskar utföra. Tyvärr, för de som sysslar med studiet av ekonomi, så består marknader av så många individer att även om varje aktör beter sig rationellt så blir summan av alla individers beteende irrationellt. För att kunna ägna oss åt ett matematiskt studium i en sådan miljö behöver vi en matematisk grund att stå på som tar hänsyn till detta. I denna uppsats använder vi oss av Itökalkyl som den grunden. Den läsare som inte är familjär med denna gren inom matematiken behöver inte känna sig avskräckt. De inledande kapitlen §3, §4 och §5 ägnar vi åt att bygga en teoretisk grund och åt att grundläggande förklara hur stokastisk analys matematiskt fungerar. Vi ägnar även en hel del utrymme åt att undersöka och förklara hur denna variant av analys skiljer sig från den klassiska matematiska analysen.

Som vi tidigare nämnt är vi som matematik-ekonomistuderande mest intresserade av de tillämpningar av matematiken som den finansiella sektorn har att erbjuda. Under 2000-talet har dualitet blivit en populär metod inom stokastisk optimering. Antagligen för att det är ett mycket kraftfullt verktyg inom vårt intresseområde finansiell kalkyl. Vi har därför valt att ägna en del av uppsatsen åt dualitet. Precis som med Itökalkylen skall den läsare som inte känner sig bevandrad inom ämnet inte avskräckas härav, då vi använder oss av samma tillvägagångssätt här. Vi lägger en god teoretisk grund innan vi går vidare till de tyngre resultaten. Dualitet är en metod för att lösa optimeringsproblem. Vi har valt att lägga det största fokuset på Lagrangedualitet, men ägnar en del kraft åt konjugatdualitet, för att få en naturlig ingång till linjär-kvadratisk kontroll.

Eftersom kopplingen mellan ämnena är relativt ny så är ämnet ännu outforskat jämfört med de flesta andra grenar inom matematiken. Vi har alltså haft fördelen att få arbeta inom ett ämnesområde som är under stark utveckling. Vi förväntar oss att stora framsteg kommer göras inom området under de närmaste åren, och ser fram emot att följa utvecklingen.

Vårt primära mål med uppsatsen har varit att få en förståelse för den involverade matematiska teorin, för att sedan, med hjälp av denna kunskap, visa på några intressanta resultat av tillämpningar av teorin. Särskilt inom dualitetsdelen har vi valt att fokusera på teorin. Det finns en stor mängd metoder för att finna exakta lösningar givet att alla nödvändiga parametrar är kända och vi uppmuntrar den läsare som är intresserad att se på lämplig litteratur inom ämnet för att finna dessa.

Till sist vill vi nämna en annan faktor som varit viktig för matematikens utveckling under senare tid, datorerna. Datorernas förmåga att utföra stora iterationer på kort tid är av stort värde för den som önskar studera stokastisk optimering. Särskilt eftersom det inte ens är tidsödande att få fram slupmässiga storheter. Eftersom vi önskat ta vara på denna resurs har vi avslutat uppsatsen med ett antal simuleringar inom några av de områden som uppsatsen berör.

Vi vill passa på att tacka vår handledare Yishao Zhou vid Matematiska institutionen på Stockholms Universitet för allt hennes jobb för vår skull.

2 Bakgrund och kort diskussion

Som vi skriver i §1 vill vi finna en matematisk modell för en slumpmässig utveckling. Särskilt kommer vi intressera oss för hur priset på riskabla tillgångar utvecklas med tiden. Vi kommer för detta syfte använda oss av standardmodellen där priset framställs som en differentialekvation

$$\frac{dS}{dt} = S(t)(r(t) + \sigma(t))$$

där $S(t)$ är priset på tillgången, $r(t)$ är den förväntade utvecklingsfaktorn och $\sigma(t)$ är ett "brus" som fångar det okända i utvecklingen. Här följer en kortare diskussion om vad vi önskar sätta in i ett matematiskt sammanhang och en idé om hur vi kan gå tillväga. Vi kan inleda med att en aning informellt skriva om uttrycket ovan som

$$dS = S(t)(r(t)dt + \sigma(t)dB_t)$$

där vi "multipliserar upp" termen dt ur den ursprungliga ekvationen (se §5.2 nedan). Vi måste dock fortfarande ta hänsyn till slumpmässigheten hos termen $\sigma(t)$ och kan därför inte låta den bero på dt , utan istället utvecklas den med avseende på en slumpmässig term dB_t som vi kommer förklara i §3 om Brownsk rörelse nedan. En differentialekvation som ovan har en lösning

$$S(t) = e^{\int_0^t r ds + \int_0^t \sigma dB_s} S(0)$$

Och vi inser att vi även måste undersöka om det går att integrera med avseende på termen dB_t om vi vill kunna utföra beräkningar och nå några resultat. Vi kommer ägna §4 åt att detta.

Innan vi går vidare och börjar utföra beräkningar måste vi definera ett sannolikhetsrum:

Definition 2.1. Ett sannolikhetsrum är en trippel (Ω, \mathbf{F}, P) där

1. Ω är en icke-tom mängd (av resultat)
2. \mathbf{F} är en mängd (av händelser)
3. P är en funktion, $P : \mathbf{F} \rightarrow [0, 1]$, som mäter sannolikheter

Vidare så gäller för dessa att

- \mathbf{F} definerar en σ -algebra på Ω , dvs

1) $\Omega \in \mathbf{F}$

2) $F \in \mathbf{F} \Rightarrow F^c \in \mathbf{F}$

- $P(\emptyset) = 0, P(\Omega) = 1$

Vi definerar också, för att undvika förvirring senare

Definition 2.2. Mängden av alla de reella talen betecknar vi med \mathfrak{R} .

Definition 2.3. De utökade reella talen, $\mathfrak{R} \cup \{-\infty, \infty\}$, är i denna uppsats betecknade \mathbf{R} .

3 Den Brownska Rörelsen

Priset på riskabla tillgångar skiftar ständigt utan att någon kan ge ett exakt svar på åt vilket håll (upp eller ner!) eller hur stort skiftet kommer vara. En matematiskt framställning av detta torde därför vara en slumpvandring i reell tid. Vi kan konstruera en sådan på följande vis:

Börja med en koordinatmatris i vilken vi låter en partikel röra sig. Varje tidssteg $n\Delta t$ tar partikeln $m\Delta x$ steg åt höger med sannolikheten $\frac{1}{2}$ eller $m\Delta x$ steg åt vänster med sannolikheten $\frac{1}{2}$. Vi låter $X(t)$ vara partikelns läge vid tiden $t = n\Delta t$. Vi kan nu definera

$$L_n = \sum_{i=1}^n X_i$$

Där X_i är oberoende stokastiska variabler sådana att

$$P(X_i = 0) = P(X_i = 1) = \frac{1}{2}, \quad i = 1, 2, 3, \dots$$

L_n är nu antalet steg åt vänster vid tiden t definerad som ovan. Observera att $E(X_i) = \frac{1}{2}$ och att $Var(X_i) = \frac{1}{4}$.

Alltså vet vi att

$$X(t) = L_n\Delta x + (n - L_n)(-\Delta x) = (2L_n - n)\Delta x$$

Vi kan nu skriva om detta som

$$X(t) = (2L_n - n)\Delta x = \left(\frac{L_n - \frac{n}{2}}{\sqrt{n/4}}\right)\sqrt{n}\Delta x \quad (3.1)$$

Låt nu $\frac{(\Delta x)^2}{\Delta t} = A$.

Så (3.1) ovan kan nu uttryckas som

$$X(t) = \left(\frac{L_n - \frac{n}{2}}{\sqrt{n/4}}\right)\sqrt{tA}$$

Låt nu $\Delta t \rightarrow 0$, $\Delta x \rightarrow 0$ och $n\Delta t \rightarrow t$.

Sannolikheten att vår partikel ska befinna sig mellan platserna a och b är nu

$$\begin{aligned} \lim_{n \rightarrow \infty} P\left(a \leq \frac{L_n - \frac{n}{2}}{\sqrt{n/4}}\sqrt{tA} \leq b\right) &= \\ &= \lim_{n \rightarrow \infty} P\left(\frac{a}{\sqrt{tA}} \leq \frac{L_n - \frac{n}{2}}{\sqrt{n/4}} \leq \frac{b}{\sqrt{tA}}\right) \end{aligned}$$

Enligt Centrala gränsvärdessatsen är denna sannolikhet lika med $\frac{1}{\sqrt{2\pi At}} \int_a^b e^{-\frac{x^2}{2At}} dx$ vilket är fördelningen för en normalfördelad variabel med väntevärde 0 och varians At . Vi låter processen börja i punkten $(0, 0)$ så $X(0) = 0$ och låter $A = 1$ och har nästan skapat Brownsk rörelse, en $\mathbf{N}(0, t)$ -fördelad variabel. Vi sammanfattar det vi gjort hittills med en definition.

Definition 3.1. Brownsk rörelse är en stokastisk process, hädanefter kallad $B(t)$ eller B_t , som är $\mathbf{N}(0, t)$ -fördelad

Vi behöver två saker till för att göra vår process användbar. För det första vill vi att varje steg skall vara oberoende av det föregående och för det andra vill vi att den ”väg” som vår process tar är kontinuerlig. Eftersom vi i resten av denna uppsats kommer hålla oss till högre dimensioner än bara en kan vi redan nu utöka vår Brownska rörelse till det n -dimensionella fallet och visa kontinuitet och stegens oberoende i detta fall. Vi använder oss av en n -vektor som innehåller n stycken oberoende en-dimensionella browniska rörelser. Istället för en normalfördelning får vi då den multi-normala fördelningen. Vi låter som ovan $B_0 = 0$. Så $E[B_t] = 0 \forall t \geq 0$.

Vi får en kovariansmatris med utseendet

$$C = \begin{pmatrix} t_1 I_n & t_1 I_n & \dots & t_1 I_n \\ t_1 I_n & t_2 I_n & \dots & t_2 I_n \\ \vdots & \vdots & & \vdots \\ t_1 I_n & t_2 I_n & \dots & t_k I_n \end{pmatrix}$$

Alltså är $E[B_t^2] = nt$ och $E[B_s B_t] = n \min(s, t)$ eller om man så vill, $E[(B_t - B_s)^2] = n(t - s)$ om $t \geq s$.

Sats 3.2. Om $s > t$ så är B_s oberoende av B_t .

Bevis. För att visa att varje steg är oberoende av det föregående använder oss av det faktum att normalfördelade variabler är oberoende av varandra om deras korrelation är lika med noll.

$$E[(B_{t_i} - B_{t_{i-1}})(B_{t_j} - B_{t_{j-1}})] = n(t_i - t_{i-1} - t_j + t_{j-1}) = 0$$

□

Det är en aning svårare att visa kontinuiteten hos rörelsens väg. Men med hjälp av *Kolmogorovs kontinuitetssats*¹ kan vi visa att det finns en version av

¹Antag att processen $X = \{X_i\}_{i \geq 0}$ uppfyller följande $\forall T \geq 0, \exists \alpha, \beta, D \geq 0$ s.a. $E[|X_t - X_s|^\alpha] \leq D \cdot |t - s|^{1+\beta}; 0 \leq s, t \leq T$. Då finns en kontinuerlig version av X .

Brownsk rörelse som är kontinuerlig genom att ta $\alpha = 4, \beta = 1, D = n(n+2)$ i nämnda sats.

Denna kontinuerliga version sammanfaller i sannolikhet ($P(B_{kont} = B_t) = 1$) med alla andra versioner av Brownsk rörelse och vi kan alltså utgå ifrån att vår rörelse är en kontinuerlig process.

I §18 kan läsas om hur man kan simulera en Brownsk rörelse med hjälp av Matlab och i §C finns några illustrationer över hur dessa kan se ut.

4 Itô's integral

Efter allt jobb med att definera den Brownska rörelsen vill vi nu få användning av den. Den kommer användas för att matematiskt uttrycka den slumpmässighet som vi måste ta hänsyn till och som vi diskuterat ovan. Det bästa sättet att göra detta är att integrera över den. Men då måste vi först definera Itô's integral. Låt som ovan B_t vara en Brownsk rörelse. Vi vill nu försöka finna vad

$$\int_S^T f(X_t, t) dB_t \quad (4.1)$$

är för något. Tyvärr kan vi inte använda oss av den "vanliga" Riemann-Stieltjes-integralen i detta fall, eftersom:

Antag att vi vill integrera (4.1) ovan och att $f(X_t, t)$ kan uttryckas som

$$\psi(t, \omega) = \sum_{j \geq 0} e_j(\omega) \cdot \chi_{[j \cdot 2^{-n}, (j+1)2^{-n}]}(t)$$

Där n är ett positivt heltal och χ är den karakteristiska funktionen. Det vore nu logiskt för oss att definera

$$\int_S^T \psi(t, \omega) dB_t(\omega) = \sum_{j \geq 0} e_j(\omega) [B_{t_{j+1}} - B_{t_j}](\omega)$$

Detta leder dock till problem, som enklast visas genom följande exempel:

$$\psi_1(t, \omega) = \sum_{j \geq 0} B_{j \cdot 2^{-n}}(\omega) \cdot \chi_{[j \cdot 2^{-n}, (j+1)2^{-n}]}(t)$$

$$\psi_2(t, \omega) = \sum_{j \geq 0} B_{(j+1) \cdot 2^{-n}}(\omega) \cdot \chi_{[j \cdot 2^{-n}, (j+1)2^{-n}]}(t)$$

Då är

$$E\left[\int_0^T \psi_1 dB_t\right] = \sum_{j \geq 0} E[B_{t_j} (B_{t_{j+1}} - B_{t_j})] = 0,$$

p.g.a. att B_t 's steg är oberoende av varandra, och

$$E\left[\int_0^T \psi_2 dB_t\right] = \sum_{j \geq 0} E[B_{t_{j+1}} (B_{t_{j+1}} - B_{t_j})] = T$$

eftersom vi får en teleskopsumma och $E[(B_t)^2] = t$.

Så ovanstående metod fungerar tyvärr inte (den ger två helt olika svar när den borde returnera samma). Istället kan vi använda oss av följande definition:

Låt A_1, A_2, \dots och A vara stokastiska variabler s.a. $E[A_n^2] < \infty$ och $E[A^2] < \infty$.

Om nu $\lim_{n \rightarrow \infty} E[(A_n - A)^2] = 0$ säger vi att $A_n \rightarrow A$ i kvadratisk medelvärde².

Vi definerar en integral m.a.p. en stokastisk process som gränsvärdet av summan (4.2) nedan som kvadratisk medelvärde.

$$\sum_{k=0}^n X(t_k)[B(t_{k+1}) - B(t_k)] \quad (4.2)$$

när $n \rightarrow \infty$, partitionen $0 = t_0 \leq t_1 \leq \dots \leq t_n = T$ blir finare och finare och där X är en stokastisk process och B_t är Brownsk rörelse. För att denna definition skall vara användbar är det viktigt att $X \in \mathbf{L}^2$, där \mathbf{L}^2 är mängden av alla progressivt mätbara funktioner. Detta innebär för allt vad vi bryr oss om egentligen att vad vi vet om X i detta nu enbart beror på vad vi observerat om X innan och inte på vad som händer i framtiden.

Exempel 1. Nu när vi har definitionen av en stokastisk integral kan vi också beräkna sådana integraler. För ett första exempel på hur detta går till väga väljer vi en stokastisk process vi väl känner till och har definerat, den browniska rörelsen, och ser efter vad $\int_0^t B_s dB_s$ blir för något. Vi inleder med att konstatera att:

$$\int_0^t B_s dB_s = \lim_{\Delta t_k \rightarrow 0} \sum_k B_{t_k} \Delta B_{t_k}$$

För enkelhets skull låter vi hädanefter B_{t_k} förkortas till B_k . Vi måste nu finna vad ovanstående summa blir. Med kunskapen om att en Riemann-integral hade returnerat $\frac{1}{2}B_k^2$ i motsvarande situation hoppas vi att samma term skall uppkomma även här. Vi utvecklar därför B_k^2 :

$$B_k^2 = \sum_k \Delta(B_k^2) \quad (4.3)$$

Om vi nu går vidare och undersöker hur $\Delta(B_k^2)$ kan uttryckas får vi:

$$\Delta(B_k^2) = B_{k+1}^2 - B_k^2$$

Som vi kan skriva om med hjälp av:

$$\Delta(B_k)^2 = (B_{k+1} - B_k)^2 = B_{k+1}^2 + B_k^2 - 2B_k B_{k+1}$$

²Fritt översatt från engelskans *mean square convergence*

$$2B_k \Delta B_k = 2B_k(B_{k+1} - B_k) = 2B_k B_{k+1} - 2B_k^2$$

$$\Delta(B_k)^2 + 2B_k \Delta B_k = B_{k+1}^2 + B_k^2 - 2B_k B_{k+1} + 2B_k B_{k+1} - 2B_k^2 = B_{k+1}^2 - B_k^2$$
 Så nu kan vi uttrycka summan (4.3) som:

$$\begin{aligned}
 B_k^2 &= \sum_k \Delta(B_k^2) = \sum_k B_{k+1}^2 - B_k^2 = \\
 &= \sum_k \Delta(B_k)^2 + 2B_k \Delta B_k = \sum_k \Delta(B_k)^2 + 2 \sum_k B_k \Delta B_k \Leftrightarrow \\
 &\Leftrightarrow \sum_k B_k \Delta B_k = \frac{1}{2}(B_k^2 - \sum_k \Delta(B_k)^2)
 \end{aligned}$$

Om vi nu låter $\Delta t_k \rightarrow 0$ i det sista steget och använder oss av ett senare resultat (5.1) får vi:

$$\int_0^t B_s dB_s = \lim_{\Delta t_k \rightarrow 0} \sum_k B_{t_k} \Delta B_{t_k} = \lim_{\Delta t_k \rightarrow 0} \frac{1}{2}(B_k^2 - \sum_k \Delta(B_k)^2) = \frac{1}{2}B_k^2 - \frac{1}{2}t$$

△

4.1 Stokastiska integraler över funktioner

Om vi nu vill kunna integrera funktioner $f(\omega, t)$ på detta vis kan vi göra som följer:

Låt ϕ vara en elementär funktion, d.v.s vara på formen

$$\phi = \sum_j e_j(\omega) \cdot \chi_{[t_j, t_{j+1})}$$

För en sådan funktion kan vi definera

$$\int \phi(t, \omega) dB_t(\omega) = \sum_j e_j(\omega) [B_{t_{j+1}} - B_{t_j}](\omega) \quad (4.4)$$

Läsaren kan kontrollera att (4.4) fungerar utmärkt för ψ_1 ovan men inte för ψ_2 , där ψ_1, ψ_2 def som ovan. Målet är nu att visa att en funktion $f(t, \omega)$ kan approximeras av en funktion $\phi(t, \omega)$ som ovan och att vi därmed kan definera

$$\int_S^T f(t, \omega) dB_t(\omega) = \lim_{n \rightarrow \infty} \int_S^T \phi_n(t, \omega) dB_t(\omega) \quad (4.5)$$

Att lyckas med detta är en omständigt process som bland annat innebär att man först approximerar den elementära funktionen ϕ_n med en begränsad

funktion $g(\cdot, \omega)$, sedan detta $g(\cdot, \cdot)$ med en funktion $h(\cdot, \cdot)$ och slutligen funktionen $f(\cdot, \cdot)$ (vårt mål från början) med en sekvens $\{h_n\}$. För alla detaljer om hur man kan göra detta hänvisar vi till [1] sid 26-29.

Ett viktigt resultat i [1] på vägen dit tycker vi kan förtjäna att omnämnas eftersom det kommer användas av oss i fortsättningen. Det är den så kallade Itô-isometrien och innebär

$$E[(\int_0^T \phi(t, \omega) dB_t(\omega))^2] = E[\int_0^T \phi(t, \omega)^2 dt]$$

Där ϕ är en elementär funktion. Beviset för detta är ganska enkelt; Sätt $\Delta B_j = B_{j+1} - B_j$. Då är

$$E[e_i e_j \Delta B_i \Delta B_j] = \begin{cases} 0 & , i \neq j \\ E[e_j^2](t_{j+1} - t_j) & , i = j \end{cases}$$

eftersom ΔB_i och ΔB_j är oberoende om $i \neq j$. Vi får

$$E[(\int_0^T \phi dB_t)^2] = \sum_{i,j} E[e_i e_j \Delta B_i \Delta B_j] = \sum_j E[e_j^2](t_{j+1} - t_j) = E[\int_0^T \phi^2 dt] \tag{4.6}$$

Detta visar att vi kan integrera över den stokastiska processen $B(\omega) \in N[0, t]$. Som vi tidigare nämnt är denna integral inte definerad som den "vanliga" Riemann-integralen, utan är en variant av Lebesgue-integralen. Vi hänvisar till [7] för teorin bakom denna integralvariant.

Vi har samlat några idéer om hur detta kan simuleras i §18, och lite tabeller över skattade värden finns i §A.

5 Itô's formel och -process

Om vi återgår till det vi från början var intresserade av, nämligen att finna en matematisk modell för ett fenomen som har en osäkerhet eller störning i sin mätning, inser vi att det inte räcker att enbart studera integraler över $B(\omega)$. Den browniska rörelsen används för att matematiskt förklara den osäkra delen i en beräkning. Vidare är det, precis som i fallet med Riemanns integral, opraktiskt att använda sig av definitionen av integralen när man vill göra beräkningar med den. Därför vill vi finna en bra väg att "blanda" integraler över $B(\omega)$ och integraler över t (eller liknande icke-stokastisk funktion). För att lyckas med ovanstående måste vi definiera Itôprocessen.

5.1 Itôprocesser

Definition 5.1. Itôprocess

En Itôprocess, eller stokastisk integral, är en stokastisk process X_t med utseendet

$$X_t = X_0 + \int_0^t f(s, \omega) ds + \int_0^t g(s, \omega) dB_s$$

Där B_t är brownisk rörelse, $P(\int_0^t g^2 ds < \infty) = 1$ och $P(\int_0^t |f| ds < \infty) = 1$ för alla $t \geq 0$.

5.2 Förenklad notation för Itôprocesser

För att mer komprimerat kunna skriva ner Itôprocesser och stokastiska ekvationer inför vi en mer kortfattat notation. Detta då beräkningar och formler annars hade blivit oöverskådligt stora och svåra att ta in. Vi låter Itôprocessen

$$X_t - X_0 = \int_0^t f(s, \omega) ds + \int_0^t g(s, \omega) dB_s$$

förkortas till

$$dX_t = f dt + g dB_t$$

Den förenklade notation har ingen egen betydelse utan är endast till för att, som namnet antyder, förenkla för läsaren och författaren. Med denna notation kan vi till exempel uttrycka resultatet (4.6) ovan som

$$(dB_t)^2 = dt \tag{5.1}$$

Ett viktigt resultat som vi redan använt och snart återkommer till.

5.3 Martingaler

För att förstå användbarheten hos Itökalkylen måste vi först introducera martingalen och se vad den har för följder och användningsområden.

Definition 5.2. Martingal

På ett sannolikhetsrum (Ω, N, P) , där $N_t \subset N$ för alla $t \geq 0$ och $\{N_t\}_{t \geq 0}$ är en expanderande σ -algebra så är:

En *martingal*, M_t , är en stokastisk process som uppfyller

$$\begin{aligned} E[|M_t|] &< \infty \\ E[M_s | N_t] &= M_t \end{aligned} \tag{5.2}$$

för alla $s \geq t \geq 0$.

Om (5.2) ovan istället skrivs

$$M_t \geq E[M_s | N_t]$$

så är M_t en supermartingal och i fallet

$$M_t \leq E[M_s | N_t]$$

så kallar vi M_t för en submartingal.

Vi har redan stött på några exempel på martingaler i denna uppsats, exempelvis är brownisk rörelse, B_t och Itös integral $\int_0^t f dB_s$ martingaler, något som vi visar i följande exempel:

Exempel 2. Brownisk rörelse är en martingal eftersom:

$$E[B_s | B_t] = E[B_s - B_t + B_t | B_t] = E[B_s - B_t | B_t] + E[B_t | B_t] = 0 + B_t = B_t$$

då vi redan tidigare konstaterat att brownisk rörelse alltid är oberoende av föregående steg och har väntevärde 0. Vidare är:

$$E\left[\int_0^t f dB_s \mid \int_0^{t-1} f dB_s\right] = 0 + E\left[\int_0^{t-1} f dB_s \mid \int_0^{t-1} f dB_s\right]$$

en martingal av samma anledning som Brownisk rörelse är det ovan och det faktum att väntevärdet av stokastiska integraler över dB_t alltid är lika med 0. \triangle

5.4 Itô's formel

Det grundläggande resultatet som Itô fann och som därför lånat hans namn till den del av matematiken som benämns Itökalkyl är Itô's formel. Den säger följande:

Sats 5.3. Itô's formel

Låt X_t vara en Itôprocess, d.v.s $dX_t = f dt + g dB_t$. Då är $Y_t = h(t, X_t)$ också en Itôprocess, och

$$dY_t = \frac{\partial h}{\partial t}(t, X_t)dt + \frac{\partial h}{\partial x}(t, X_t)dX_t + \frac{1}{2} \frac{\partial^2 h}{\partial x^2}(t, X_t) \cdot (dX_t)^2$$

Där vi kan beräkna den sista termen $(dX_t)^2 = (dX_t) \cdot (dX_t)$ enligt regeln $dtdt = dtdB_t = dB_t dt = 0$ och resultatet (5.1).

Bevis. Tag Itôprocessen $h(t, X_t)$ och Taylorutveckla den, vi får

$$\begin{aligned} h(t, X_t) &= h(0, X_0) + \sum_i \frac{\partial h}{\partial t} \Delta t_i + \sum_i \frac{\partial h}{\partial x} \Delta X_i + \\ &\frac{1}{2} \sum_i \frac{\partial^2 h}{\partial t^2} (\Delta t_i)^2 + \sum_i \frac{\partial^2 h}{\partial t \partial x} (\Delta t_i)(\Delta X_i) + \frac{1}{2} \sum_i \frac{\partial^2 h}{\partial x^2} (\Delta X_i)^2 + \sum_i R_i \end{aligned}$$

Där R_i är en restterm. Vi låter $\Delta t_i \rightarrow 0$ och får:

$$\sum_i \frac{\partial h}{\partial t} \Delta t_i = \sum_i \frac{\partial h}{\partial t}(t_i, X_{t_i}) \Delta t_i \longrightarrow \int_0^t \frac{\partial h}{\partial t}(s, X_s) dt \quad (5.3)$$

$$\sum_i \frac{\partial h}{\partial x} \Delta X_i = \sum_i \frac{\partial h}{\partial x}(t_i, X_{t_i}) \Delta X_i \longrightarrow \int_0^t \frac{\partial h}{\partial x}(s, X_s) dX_s \quad (5.4)$$

Vidare så gäller att:

$$\begin{aligned} \frac{1}{2} \sum_i \frac{\partial^2 h}{\partial t^2} (\Delta t_i)^2 &\longrightarrow 0 \\ \sum_i \frac{\partial^2 h}{\partial t \partial x} (\Delta t_i)(\Delta X_i) &\longrightarrow 0 \\ \sum_i R_i &\longrightarrow 0 \end{aligned}$$

Och vi behöver nu bara klara av den näst sista termen. Vi skriver om den som:

$$\frac{1}{2} \sum_i \frac{\partial^2 h}{\partial x^2} (\Delta X_i)^2 = \underbrace{\sum_i \frac{\partial^2 h}{\partial x^2} f_i^2 (\Delta t_i)^2}_{\rightarrow 0, \Delta t_i \rightarrow 0} + 2 \underbrace{\sum_i \frac{\partial^2 h}{\partial x^2} f_i g_i (\Delta t_i)(\Delta B_i)}_{\rightarrow 0, \Delta t_i \rightarrow 0} + \sum_i \frac{\partial^2 h}{\partial x^2} g_i^2 (\Delta B_i)^2$$

Om vi återigen återoppar resultatet (5.1) kan vi skriva om den resterande termen och se på dess gränsvärde:

$$\sum_i \frac{\partial^2 h}{\partial x^2} g_i^2(\Delta B_i)^2 = \sum_i \frac{\partial^2 h}{\partial x^2} g_i^2(\Delta t_i) \longrightarrow \int_0^t \frac{\partial^2 h}{\partial x^2} (g_i(s, X_s))^2 ds \quad (5.5)$$

Alltså kan vi använda (5.3), (5.4) och (5.5) och uttrycka Itōs formel som:

$$dY_t = \frac{\partial h}{\partial t}(t, X_t)dt + \frac{\partial h}{\partial x}(t, X_t)dX_t + \frac{1}{2} \frac{\partial^2 h}{\partial x^2}(t, X_t)dt$$

Där X_t och $Y_t = h(t, X_t)$ är definierade som ovan i Sats 5.3. \square

Om vi drar oss till minnes hur lång tid det tog att beräkna Exempel 1 när vi använde oss av definitionen av stokastiska integraler kan vi nu med hjälp av Itōs formel göra samma beräkning mycket enklare och snabbare:

Exempel 3. Precis som i Exempel 1 utgår vi från $\int_0^t B_s dB_s$. Eftersom vi från klassisk analys förväntar oss att svaret ska bli $\frac{1}{2}B_t^2$ väljer vi funktionen $h(t, x) = \frac{1}{2}x^2$. Naturligtvis använder vi B_t som vårt X_t i Sats 5.3. Alltså är vårt $Y_t = h(t, B_t) = \frac{1}{2}B_t^2$ och Itōs formel ger:

$$dY_t = \frac{\partial h}{\partial t}dt + \frac{\partial h}{\partial x}dB_t + \frac{1}{2} \frac{\partial^2 h}{\partial x^2}(dB_t)^2 = 0dt + B_t dB_t + \frac{1}{2} \cdot 1 \cdot (dB_t)^2 = B_t dB_t + \frac{1}{2}dt$$

Så alltså gäller

$$B_t dB_t = d\left(\frac{1}{2}B_t^2\right) - \frac{1}{2}dt \Leftrightarrow \int_0^t B_s dB_s = \frac{1}{2}B_t^2 - \frac{1}{2}t$$

precis som i Exempel 1. \triangle

5.5 Itōs formel i n-dimensionella fallet

Att utöka Sats 5.3 till att innefatta fler dimensioner medför inga som helst problem. Som i fallet med en dimension utgår vi från den (n-dimensionella) Itōprocessen $dX_t = f dt + \underline{g} dB_t$ där:

$$X_t = \begin{pmatrix} X_1(t) \\ X_2(t) \\ \vdots \\ X_n(t) \end{pmatrix}, f = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix}, \underline{g} = \begin{pmatrix} v_{11} & \dots & v_{1m} \\ \vdots & & \vdots \\ v_{n1} & \dots & v_{nm} \end{pmatrix}, dB_t = \begin{pmatrix} dB_1 \\ \vdots \\ dB_m \end{pmatrix}$$

Nu kan vi formulera

Sats 5.4. Den n -dimensionella Itô's formel

Låt X_t vara en n -dimensionell Itôprocess definerad som ovan s.a. även $Y_t = h(t, X_t)$ är en Itôprocess. Då gäller:

$$dY_k = \frac{\partial h_k}{\partial t}(t, X)dt + \sum_i \frac{\partial h_k}{\partial x_i}(t, X)dX_i + \sum_{i,j} \frac{\partial^2 h_k}{\partial x_i \partial x_j}(t, X)dX_i dX_j$$

Där $dt dt = dt dB_i = dB_i dt = 0$ och $(dB_i)(dB_j) = \delta_{ij} dt$.

Beviset är analogt med beviset för det endimensionella fallet, förutom att alla beräkningar blir i matrisform, och vi låter bli att gå in på det här utan hänvisar till beviset för Sats 5.3.

5.6 Några beräkningstekniska resultat

Som vi diskuterat ovan så kan man inte lita på att kända resultat från den deterministiska integralkalkylen gäller för stokastiska integraler. Dessutom vill vi än en gång passa på att poängtera några detaljer. Det första vi vill nämna är partiell integration, i "vanlig" integralkalkyl gäller följande resultat:

$$\int f dg = fg - \int g df$$

Beviset för detta är en ganska simpel kombination av kedjeregeln och analysens huvudsats:

$$fg = \int d(fg) = \int f dg + \int g df \Leftrightarrow \int f dg = fg - \int g df$$

Om vi nu istället är intresserade av vad som händer i det stokastiska fallet, dvs när vi har integralen

$$\int_0^t X_s dY_s$$

där X_t och Y_t är Itôprocesser, kan vi finna detta resultat genom Itô's formel.

Sats 5.5. Formeln för stokastisk partiell integration

I stokastisk integralkalkyl gäller att

$$\int_0^t X_s dY_s = [X_s Y_s]_0^t - \int_0^t Y_s dX_s - \int_0^t dY_s dX_s$$

Genom Sats 5.4 får vi ett enkelt bevis

Bevis. Låt X_t och Y_t vara Itôprocesser. Sätt $Z_t = h(X_t, Y_t, t) = xy$ och applicera Itôs (2-dimensionella) formel på Z_t . Vi får

$$\begin{aligned}
d(X_t Y_t) &= dZ_t = \frac{\partial h}{\partial t}(X_t, Y_t, t)dt + \frac{\partial h}{\partial x}(X_t, Y_t, t)dX_t + \\
&+ \frac{\partial h}{\partial y}(X_t, Y_t, t)dY_t + \frac{1}{2} \sum \frac{\partial^2 h}{\partial x \partial y}(X_t, Y_t, t)dX_t dY_t = \\
&= 0 + Y_t dX_t + X_t dY_t + \frac{1}{2}(dY_t dX_t + dX_t dY_t) = \\
&= Y_t dX_t + X_t dY_t + dY_t dX_t = d(X_t Y_t) \iff \\
&\iff X_t dY_t = d(X_t Y_t) - Y_t dX_t - dY_t dX_t \iff \\
&\iff \int_0^t X_s dY_s = [X_t Y_t]_0^t - \int_0^t Y_s dX_s - \int_0^t dX_s \cdot dY_s
\end{aligned}$$

□

Vi har tidigare använt oss av att väntevärdet av (se exempelvis §5.3 om Martingaler) en stokastisk integral m.a.p. $B(\omega)$ är noll, det kan vara på sin plats att visa att detta resultat faktiskt gäller:

Sats 5.6. Väntevärdet av Itôs integral

Låt B_t vara brownsk rörelse och $f(s, \cdot)$ någon funktion s.a.

$$E\left[\int_0^T f(s, \cdot)dB_s\right]$$

är väntevärdet av en Itôintegral. Då gäller att detta väntevärde är lika med noll.

Bevis. Kom ihåg att $E[B_t] = 0$ och att

$$\sum_{n=0}^{N-1} f(t_n, \omega) \Delta B_n \longrightarrow \int_0^T f(s, \omega) dB_s$$

för en partitionering $0 = t_0 < t_1 < \dots < t_{N-1} < t_N = T$.

Alltså gäller att:

$$E\left[\sum_{n=0}^{N-1} f(t_n, \omega) \Delta B_n\right] = \sum_{n=0}^{N-1} E[f(t_n, \omega)]E[\Delta B_n] = 0$$

Och vi kan sluta oss till att

$$0 = E\left[\sum_n f(t_n, \omega) \Delta B_n\right] \rightarrow E\left[\int_0^T f(s, \omega) dB_s\right] = 0$$

□

Det kan också vara av intresse att kontrollera hur väl Itōs integral följer de resultat som vi är vana vid från Riemannsk integralkalkyl.

Sats 5.7. *Följande gäller för integraler över dB_s :*

- 1) $\int_0^T cf(s, \omega)dB_s = c \int_0^T f(s, \omega)dB_s$ där $c \in \mathfrak{R}$
- 2) $\int_0^T f_1 + f_2dB_s = \int_0^T f_1dB_s + \int_0^T f_2dB_s$
- 3) Resultaten 1) och 2) gäller samtidigt och tillsammans
- 4) $E[(\int f(s, \omega)dB_s)(\int g(s, \omega)dB_s)] = \int E[(fg)(s, \omega)]ds$

Som alla fyra går att visa med hjälp av definitionen av Itōs integral, vi visar 3) och 4).

Bevis. För att visa 3), antag att $c_1, c_2 \in \mathfrak{R}$ och att f_1 och f_2 är två Itōintegrerbara funktioner. Då gäller att $\int_0^T (c_1f_1(s, \omega) + c_2f_2(s, \omega))dB_s$ är gränsvärdet av:

$$\lim_{\Delta t_n \rightarrow 0} \sum_{n=0}^{N-1} [c_1f_1(t_n, \omega) + c_2f_2(t_n, \omega)]\Delta B_n$$

som vi kan skriva om som

$$\begin{aligned} & \sum_{n=0}^{N-1} [c_1f_1(t_n, \omega) + c_2f_2(t_n, \omega)]\Delta B_n = \sum_n c_1f_1(\cdot)\Delta B_n + \sum_n c_2f_2(\cdot)\Delta B_n = \\ & = c_1 \sum_n f_1\Delta B_n + c_2 \sum_n f_2\Delta B_n \longrightarrow c_1 \int_0^T f_1(s, \omega)dB_s + c_2 \int_0^T f_2(s, \omega)dB_s \end{aligned}$$

Vilket visar 3), 2) och 1).

Vi visar 4) genom att notera att

$$\left(\sum_{n=0}^{N-1} f(t_n, \cdot)\Delta B_n\right)\left(\sum_{n=0}^{N-1} g(t_n, \cdot)\Delta B_n\right) \rightarrow \left(\int_0^T f(s, \cdot)dB_s\right)\left(\int_0^T g(s, \cdot)dB_s\right)$$

tillsammans med att samma produkt av dessa summor också kan uttryckas som

$$\begin{aligned} & \left(\sum_{n=0}^{N-1} f(t_n, \cdot)\Delta B_n\right)\left(\sum_{n=0}^{N-1} g(t_n, \cdot)\Delta B_n\right) = \\ & = (f_0\Delta B_0)(g_0\Delta B_0) + (f_1\Delta B_1)(g_1\Delta B_1) + \dots = \end{aligned}$$

$$= f_0 g_0 (\Delta B_0)^2 + f_1 g_1 (\Delta B_1)^2 + \dots = \sum_{n=0}^{N-1} f_n g_n (\Delta B_n)^2$$

Där vi använt beteckningen f_u för $f(t_u, \omega)$ där $u \in [0, N - 1]$.
Om vi nu tar väntevärdet av ovanstående ser vi att:

$$\begin{aligned} & E\left[\left(\sum_{n=0}^{N-1} f(t_n, \cdot) \Delta B_n\right) \left(\sum_{n=0}^{N-1} g(t_n, \cdot) \Delta B_n\right)\right] = \\ &= E\left[\sum_{n=0}^{N-1} f_n g_n (\Delta B_n)^2\right] = \sum_{n=0}^{N-1} E[f_n g_n] \Delta t_n \rightarrow \\ &\rightarrow \int_0^T E[f(s, \omega) g(s, \omega)] dt = \int_0^T E[(fg)(s, \omega)] ds \end{aligned}$$

Där vi ovan återigen använt resultatet (5.1) från Itōisometriken. Detta visar 4). □

6 Stokastiska differentialekvationer

Vi drar oss till minnes att vi i §2 använde oss av en differentialekvation med stokastiska komponenter för att uttrycka priset på en riskabel tillgång. Vi ställer oss nu naturligt frågan om det går att ställa upp en sådan differentialekvation samt om den går att lösa. Svaret är förstås att ja, det går att formulera och lösa sådana problem. I det här kapitlet ska vi visa hur det går till.

6.1 Uppställning av stokastisk differential ekvation, SDE

I §2 använde vi oss av uttrycket

$$\frac{dS}{dt} = S(t)(r(t) + \sigma(t))$$

för vår differentialekvation. Med hjälp av tidigare resultat kan vi nu skriva om detta uttryck. Slumpmässigheten hos $\sigma(t)$ beskriver vi nu som att den beror på en stokastisk process som vi kan kalla C_t . Vi kan också uttrycka $r(t)$ och $\sigma(t)$ enligt notationen från §4. Därmed får uttrycket formen:

$$\begin{aligned} \frac{dS_t}{dt} = r_t S_t + S_t \sigma(t, C_t) &\Leftrightarrow \frac{dS_t}{dt} = f(t, S_t) + h(t, S_t) C_t \\ &\Leftrightarrow dS_t = f(t, S_t) dt + h(t, S_t) dB_t \end{aligned} \quad (6.1)$$

Där vi låtit $C_t = \frac{dB_t}{dt}$ vara bruset. Diff-ekvationen är alltså en Itôprocess. Vi ser att vi kan skriva om (6.1) genom att dela med S_t i båda led:

$$\frac{dS_t}{S_t} = f_* dt + h_* dB_t$$

Denna framställning är viktig för ekonomiska applikationer eftersom den mäter förändring i pris på en riskabel tillgång om vi låter S_t vara priset och $f_* = \mu, h_* = \sigma$ vara konstanter som mäter drift och volatilitet. Framställningen har också en lösning som vi redan talat om i §2. Vi ska utreda hur vi löser en SDE och särskilt visa att lösningen i §2 är korrekt med hjälp av Itôs formel (Sats 5.3) nedan. Vi inleder med en definition:

Definition 6.1. SDE

En stokastisk process, X_t , löser

$$dX_t = f(X_t, t) dt + h(X_t, t) dB_t, \quad t \geq 0, \quad X_0 = x_0$$

Om

$$\forall t \geq 0 : X_t = x_0 + \int_0^t f(X_s, s) ds + \int_0^t h(X_s, s) dB_s$$

6.2 Lösning av enkla SDEs

En mekanisk väg att lösa en SDE som har utseendet:

$$dX_t = f(X_t, t)dt + h(X_t, t)dB_t = a^* f^*(X_t, t)dt + a^* h^*(X_t, t)dB_t \quad (6.2)$$

Där a^* är ett element, innehållandes X_t, t eller båda som finns i både f och h , går till som följer

Steg 1

Dividera båda led med a^* : (6.2) $\Rightarrow \frac{dX_t}{a^*} = f^*(X_t, t)dt + h^*(X_t, t)dB_t$

Exempel: $dX_t = X_t r_t dt + X_t \sigma_t dB_t \Rightarrow \frac{dX_t}{X_t} = r_t dt + \sigma_t dB_t$

Steg 2

Använd Itô's formel (Sats 5.3) för att finna ett annat uttryck för integralen i V.L. ovan.

$$\begin{aligned} \text{Exempel: } g(x, t) = \ln x &\Rightarrow d(\ln X_t) = 0dt + \frac{dX_t}{X_t} + \frac{-1}{2X_t^2}(dX_t)^2 = \\ &= \frac{dX_t}{X_t} - \frac{\sigma^2}{2}dt \Leftrightarrow \frac{dX_t}{X_t} = d(\ln X_t) + \frac{\sigma^2}{2}dt \end{aligned}$$

Steg 3

Sätt uttrycken från Steg 1 och Steg 2 lika med varandra och lös ut X_t .

Exempel: $r_t dt + \sigma_t dB_t = d(\ln X_t) + \frac{\sigma_t^2}{2}dt \Leftrightarrow d(\ln X_t) = (r_t - \frac{\sigma_t^2}{2})dt + \sigma_t dB_t \Rightarrow^3$

$\Rightarrow \ln(\frac{X_t}{X_0}) = \int_0^t (r_s - \frac{\sigma_s^2}{2})ds + \int_0^t \sigma_s dB_s \Leftrightarrow X_t = X_0 \exp(\int_0^t (r_s - \frac{\sigma_s^2}{2})ds + \int_0^t \sigma_s dB_s)$

Vårt exempel visar att vi hade rätt då vi i §2 påstod en lösning för den SDE vi där ställde upp, om vi låter $r = r_t - \sigma_t^2/2$ och $\sigma = \sigma_t$. Observera dock att med r och σ som konstanter kan vi förenkla lösningen till:

$$S_t = S_0 e^{(r_t - \frac{\sigma_t^2}{2}) + \sigma_t B_t} = S_0 e^{rt + \sigma B_t}$$

Vilket också leder oss till:

$$E[S_t] = E[S_0 e^{rt + \sigma B_t}] \underbrace{=}_{E[B_t]=0} S_0 e^{rt}$$

Som visar att väntevärdet av vår tillgångs värde stämmer överens med det deterministiska fallet, vilket är logiskt ur både ekonomisk och matematisk synpunkt.

6.3 Lösning av "svårare" SDEs

Det är enkelt att komma på två sätt att konstruera "svårare" SDEs, det första är att tänka sig att vi har fler ekvationer än en, och att lösningen alltså måste

³ $d(\ln X_t) = \ln X_t - \ln X_0 = \ln(\frac{X_t}{X_0})$

ske i matrisform:

$$dX_t = \begin{pmatrix} X_{1t} \\ X_{2t} \\ \dots \\ X_{nt} \end{pmatrix}$$

Samt det andra fallet då vi har en SDE av högre grad än ett, nedan ser vi exemplet med andra gradens SDE:

$$\frac{d}{dt} \frac{dX_t}{dt} = d(d(X_t))$$

Dessa fall är egentligen samma problem, ty antag att vi har en SDE av grad $(n - 1)$ (nedan med notation $X^{(n)} = d(d(\underbrace{\dots}_{n-2st}(dX)))$). Då kan vi göra en

variabelsubstitution som nedan:

$$\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \dots \\ Y_n \end{pmatrix} = \begin{pmatrix} X \\ X^{(1)} \\ X^{(2)} \\ \dots \\ X^{(n-1)} \end{pmatrix}$$

Och vi är tillbaka i det första fallet med en SDE i flera dimensioner. Vi visar ett exempel på hur detta kan se ut och hur det löses

Exempel 4. Låt P vara en stokastisk process och B_t en Brownsk rörelse. Studera SDE:n:

$$AP^{(2)} + DP^{(1)} + \frac{1}{C}P = \alpha_t + \beta_t B_t$$

Vi introducerar enligt proceduren ovan

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} P \\ P^{(1)} \end{pmatrix}$$

Och ser att

$$\begin{aligned} X_1^{(1)} &= X_2 \\ X_2^{(1)} &= \frac{1}{A}(\alpha_t + \beta_t B_t - DX_2 - \frac{1}{C}X_1) \end{aligned} \tag{6.3}$$

Om vi nu ställer upp matriserna

$$L = \begin{pmatrix} 1 & 0 \\ -1/C & -D \end{pmatrix}, M = \begin{pmatrix} 0 \\ \alpha_t \end{pmatrix}, N = \begin{pmatrix} 0 \\ \beta \end{pmatrix}, dX_t = \begin{pmatrix} dX_1 \\ dX_2 \end{pmatrix}$$

Kan vi skriva (6.3) i matrisform

$$dX_t = \frac{1}{A}[(LX_t + M)dt + NdB_t]$$

En klassisk lösningsväg i det icke-stokastiska fallet är att multiplicera båda led med den integrerande faktorn e^{-Lt} . Det fungerar även här i den stokastiska världen, när vi tar hjälp av:

Definition 6.2. För en $n \times n$ matris Γ är

$$e^\Gamma = \sum_{n=0}^{\infty} \frac{1}{n!} \Gamma^n$$

Alltså skriver vi (6.3) i matrisform och använder Definition 6.2 och får

$$e^{-Lt} AdX_t - e^{-Lt} LX_t dt = e^{-Lt} M dt + e^{-Lt} N dB_t \quad (6.4)$$

Vi använder sen, precis som i envariabelsfallet, Itös formel (men i två dimensioner) på (6.4):s vänsterled och får:

$$d(e^{-Lt} X_t) = (-L)e^{-Lt} X_t dt + e^{-Lt} dX_t$$

Som vi sedan sätter in i grundekvationen och får fram en lösning. \triangle

Beviset för att det går att lösa en SDE är i mångt och mycket identiskt med beviset för att det går att lösa en ordinär differential ekvation (ODE) och utelämnas därför. Däremot så kan följande sats vara intressant

Sats 6.3. Antag $T > 0$ och att $b : [0, T] \times \mathfrak{R}^n \rightarrow \mathfrak{R}^n$, $\sigma : [0, T] \times \mathfrak{R}^n \rightarrow \mathfrak{R}^{n \times m}$ är mätbara funktioner som uppfyller:

$$|b(t, x) - b(t, y)| + |\sigma(t, x) - \sigma(t, y)| \leq L |x - y|; \quad x, y \in \mathfrak{R}, t \in [0, T]$$

för någon konstant L .⁴ Då är den process X_t som löser SDE:n

$$dX_t = b(t, X_t) dt + \sigma(t, X_t) dB_t; \quad 0 \leq t \leq T, \quad X_0 = Y$$

unik.

För att kunna bevisa denna sats behöver vi ta hjälp av:

Lemma 6.4. Grönvalls olikhet

Givet att en positiv funktion $v(t)$ uppfyller

$$v(t) \leq C + A \int_0^t v(s) ds$$

för några konstanter A och C gäller att:

$$v(t) \leq C e^{At}$$

⁴För Lipschitz, olikheten är mest känd som Lipschitz villkor

Bevis.

$$\begin{aligned}
w(t) &:= \int_0^t v(s)ds \Rightarrow w'(t) = v(t) \leq C + A \int_0^t v(s)ds = C + Aw(t) \\
f(t) &:= w(t)e^{-At} \Rightarrow f'(t) = w'(t)e^{-At} - Aw(t)e^{-At} \leq e^{-At}(C + Aw(t) - Aw(t)) = Ce^{-At} \Rightarrow \\
&\Rightarrow f(t) \leq C \int_0^t e^{-As} ds = \frac{-C}{A}(e^{-At} - 1) = \frac{C}{A}(1 - e^{-At}) \\
f(t) &= w(t)e^{-At} \leq \frac{C}{A}(1 - e^{-At}) \Leftrightarrow w(t) \leq \frac{C}{A}e^{At} - \frac{C}{A} = \frac{C}{A}(e^{At} - 1) \\
v(t) &= w'(t) \leq C + Aw(t) \leq C + A\left(\frac{C}{A}(e^{At} - 1)\right) = C + Ce^{At} - C = Ce^{At}
\end{aligned}$$

□

Och nu kan vi ge oss på beviset för Sats 6.3:

Bevis. Antag att $X_1(t, \omega) = X_t$ och $X_2(t, \omega) = \hat{X}_t$ båda löser SDE:n ovan, samt att $X_0 = Y, \hat{X}_0 = \hat{Y}$. Sätt $a(s, \omega) = b(s, X_s) - b(s, \hat{X}_s)$ och $\rho(s, \omega) = \sigma(s, X_s) - \sigma(s, \hat{X}_s)$. Då gäller:

$$\begin{aligned}
E[|X - \hat{X}|^2] &= E[(Y - \hat{Y} + \int_0^t ads + \int_0^t \rho dB_s)^2] \leq \\
&\leq 3E[|Y - \hat{Y}|^2] + 3E[(\int_0^t ads)^2] + 3E[(\int_0^t \rho dB_s)^2] \leq \\
&\leq 3E[|Y - \hat{Y}|^2] + 3tE[\int_0^t a^2 ds] + 3E[\int_0^t \rho^2 ds] \leq \\
&\leq 3E[|Y - \hat{Y}|^2] + 3(1+t)L^2 \int_0^t E[|X - \hat{X}|^2] ds \Rightarrow \\
&\Rightarrow E[|X_t - \hat{X}_t|^2] \leq F + A \int_0^t E[|X_s - \hat{X}_s|^2] ds
\end{aligned}$$

Där $F = 3E[|Y - \hat{Y}|^2]$ och $A = 3(1+t)L^2$. Låt nu $v(t) = E[|X_t - \hat{X}_t|^2]$ och applicera Lemma 6.4:

$$v(t) \leq F + A \int_0^t v(s)ds \Rightarrow v(t) \leq Fe^{At}$$

Antag nu att $Y = \hat{Y} \Rightarrow F = 0 \Rightarrow v(t) = 0 \forall t \geq 0$. Alltså gäller att

$$P[(X_t - \hat{X}_t) = 0] = 1 \Rightarrow P[(X_1 - X_2) = 0] = 1$$

□

7 Inledning till dualitet

Inom matematiken så talar man om linjära och icke-linjära programmeringsproblem. Där linjära programmeringsproblem (LP-problem) består av en linjär målfunktion som ska optimeras under ett antal linjära bivillkor som består av likheter och/eller olikheter. De icke-linjära programmeringsproblemen (ILP-problem) består av att lösa ett system av likheter och olikheter över en mängd okända variabler, där en målfunktion ska optimeras. Varken bivillkoren eller målfunktionen är linjära. Givet ett icke-linjärt programmeringsproblem finns det alltid ett annat ILP-problem nära knutet till det första problemet. Det första problemet kallas för det *primala problemet* och det senare för det *duala problemet*. Vi ska i detta avsnitt beskriva de olika egenskaper som det duala problemet har. Man använder dessa egenskaper för att ta fram generella metoder för att lösa de primala och duala problemen.

Betrakta följande *icke-linjära programmeringsproblem*:

$$\begin{array}{lll} \text{Minimera} & f(\mathbf{x}) & \\ \\ \text{Då} & g_i(\mathbf{x}) \leq 0 & \text{för } i = 1, 2, \dots, m \\ & h_i(\mathbf{x}) = 0 & \text{för } i = 1, 2, \dots, k \\ & \mathbf{x} \in X & \end{array}$$

där f , g_1, \dots, g_m och h_1, \dots, h_k är funktioner definierade på \mathbf{R}^n , X är en delmängd av \mathbf{R}^n och \mathbf{x} är en vektor av n element x_1, \dots, x_n . Problemet ovan går ut på att finna variablerna x_1, \dots, x_n som uppfyller bivillkoren och som samtidigt minimerar funktionen $f(x)$.

Man brukar benämna f som *målfunktionen*, en vektor \mathbf{x} som uppfyller alla villkor sägs vara en *möjlig lösning* till problemet. Den icke-linjära programmeringen är att finna en möjlig punkt $\bar{\mathbf{x}}$ så att $f(\mathbf{x}) \geq f(\bar{\mathbf{x}})$ för varje möjlig punkt \mathbf{x} , där $\bar{\mathbf{x}}$ är den *optimala lösningen* till problemet.

Man kan givetvis också ha ett icke-linjära problem där problemet är att maximera målfunktionen och bivillkoren skrivs på formen $g_i(\mathbf{x}) \geq 0$ för $i = 1, \dots, m$.

8 Lagrangedualitet

Om vi har följande icke-linjära programmeringsproblem (P), det så kallade primala problemet.

Primala problemet (P):

$$\begin{array}{ll} \text{Minimera} & f(\mathbf{x}) \\ \\ \text{Då} & g_i(\mathbf{x}) \leq 0 \quad \text{för } i = 1, 2, \dots, m \\ & h_i(\mathbf{x}) = 0 \quad \text{för } i = 1, 2, \dots, k \\ & \mathbf{x} \in X \end{array}$$

Så är Lagranges duala problem (D) följande:

Lagranges duala problem (D):

$$\begin{array}{ll} \text{Maximera} & L(\mathbf{u}, \mathbf{v}) \\ \\ \text{Då} & \mathbf{u} \geq \mathbf{0} \end{array}$$

Där

$$L(\mathbf{u}, \mathbf{v}) = \inf \left\{ f(\mathbf{x}) + \sum_{i=1}^m u_i g_i(\mathbf{x}) + \sum_{i=1}^k v_i h_i(\mathbf{x}) : \mathbf{x} \in X \right\}$$

är *Lagranges duala funktion* och där u_i och v_i refereras till som *Lagrange-multiplikatorer*, uppkallade efter den kände matematikern Joseph Louis Lagrange. Vi kommer för att få en mer överskådlig text att använda oss av vektornotationen, dvs $L(\mathbf{u}, \mathbf{v}) = \inf \{ f(\mathbf{x}) + \mathbf{u}^t \mathbf{g}(\mathbf{x}) + \mathbf{v}^t \mathbf{h}(\mathbf{x}) : \mathbf{x} \in X \}$, där $f : \mathbf{R}^n \rightarrow \mathbf{R}^1$, $\mathbf{g} : \mathbf{R}^n \rightarrow \mathbf{R}^m$ är en vektor vars i :te komponent är g_i och $\mathbf{h} : \mathbf{R}^n \rightarrow \mathbf{R}^k$ är en vektor vars i :te komponent är h_i .

Notera att funktionen $L(\mathbf{u}, \mathbf{v})$ kan anta värdet $-\infty$ för någon vektor (\mathbf{u}, \mathbf{v}) . Multiplikatorn u_i som är associerad med olikheten $g_i(\mathbf{x}) \leq 0$ är icke-negativ, den har alltså teckenrestriktioner, däremot har multiplikatorn v_i som är associerad med likheten $h_i(\mathbf{x}) = 0$ inga restriktioner när det gäller tecken.

Då det duala problemet består av att maximera infimum (dvs den största undre begränsningen) av funktionen

$$f(\mathbf{x}) + \sum_{i=1}^m u_i g_i(\mathbf{x}) + \sum_{i=1}^k v_i h_i(\mathbf{x}) : \mathbf{x} \in X$$

kallas det ibland för det *max-min duala problemet*. Man bör skriva $\sup \{L(\mathbf{u}, \mathbf{v}) : \mathbf{u} \geq 0\}$ hellre än att skriva $\max \{L(\mathbf{u}, \mathbf{v}) : \mathbf{u} \geq 0\}$. Då maximum inte alltid existerar.

Exempel 5.

$$\text{Minimera } z = \sum_{i=1} c_j x_j$$

$$\text{Då } x_j \geq 0$$

$$\iff$$

$$\min c^t x$$

$$\text{då } \begin{aligned} b - Ax &\leq 0 \\ x &\geq 0 \end{aligned}$$

Den duala funktionen är då

$$\begin{aligned} L(u) &= \min\{c^t x + u^t(b - Ax) \mid x \geq 0\} \\ &= \min\{(c^t - u^t A)x + u^t b \mid x \geq 0\} \\ &= \min\{\sum_{j=1}^n (c^t - u^t A)_j x_j + u^t b \mid x_j \geq 0\} \end{aligned}$$

Om något $(c^t - u^t A)_j < 0$ så kan motsvarande x_j öka obegränsat och minvärdet blir $-\infty$. Om istället alla $(c^t - u^t A)_j \geq 0$ så lönar det sig att bäst att sätta varje $x_j = 0$, då blir:

$$L(u) = \begin{cases} u^t b & \text{då } c^t - u^t A \geq 0 \\ -\infty & \text{annars} \end{cases}$$

Eftersom vi vill maximera $L(u)$ så är inte $-\infty$ av intresse. Vidare har vi en relaxering bara för $u \geq 0 \implies$

$$\begin{aligned} \max L(u) & & \max & u^t b \\ \text{då } u &\geq 0 & \iff & \text{då } \begin{aligned} c^t - u^t A &\geq 0 \\ u &\geq 0 \end{aligned} \end{aligned}$$

△

8.1 Svag dualitet

I detta avsnitt kommer vi att visa att målfunktionens värde för någon nåbar lösning till det duala problemet ger en undre gräns till målfunktionens värde för någon nåbar lösning till det primala problemet. Detta refereras till som den *Svaga dualitetssatsen*. Några andra viktiga resultat följer som följsatser.

Sats 8.1. Svaga dualitetssatsen

Låt \mathbf{x} vara en nåbar lösning till problem P , alltså $\mathbf{x} \in X$, $\mathbf{g}(\mathbf{x}) \leq \mathbf{0}$ och $\mathbf{h}(\mathbf{x}) = \mathbf{0}$. Låt också (\mathbf{u}, \mathbf{v}) vara en nåbar lösning till problem D , alltså $\mathbf{u} \geq \mathbf{0}$. Då är

$$f(\mathbf{x}) \geq L(\mathbf{u}, \mathbf{v})$$

Bevis. Genom definitionen av L och eftersom $\mathbf{x} \in X$ så har vi att för något $\mathbf{y} \in X$

$$\begin{aligned} L(\mathbf{u}, \mathbf{v}) &= \inf \{f(\mathbf{y}) + \mathbf{u}^t \mathbf{g}(\mathbf{y}) + \mathbf{v}^t \mathbf{h}(\mathbf{y}) : \mathbf{y} \in X\} \\ &\leq f(\mathbf{x}) + \mathbf{u}^t \mathbf{g}(\mathbf{x}) + \mathbf{v}^t \mathbf{h}(\mathbf{x}) \leq f(\mathbf{x}) \end{aligned}$$

eftersom $\mathbf{u} \geq \mathbf{0}$, $\mathbf{g}(\mathbf{x}) \leq \mathbf{0}$ och $\mathbf{h}(\mathbf{x}) = \mathbf{0}$. □

Från denna sats ser man ganska enkelt att följande hjälpsatser gäller, vi lämnar bevisen till läsaren att själv fundera över.

Följsats 8.2.

$$\inf \{f(\mathbf{x}) : \mathbf{x} \in X, \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \mathbf{h}(\mathbf{x}) = \mathbf{0}\} \geq \sup \{L(\mathbf{u}, \mathbf{v}) : \mathbf{u} \geq \mathbf{0}\}$$

Följsats 8.3.

Om $f(\bar{\mathbf{x}}) = L(\bar{\mathbf{u}}, \bar{\mathbf{v}})$ där $\bar{\mathbf{u}} \geq \mathbf{0}$ och $\bar{\mathbf{x}} \in \{\mathbf{x} \in X : \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \mathbf{h}(\mathbf{x}) = \mathbf{0}\}$ då löser $\bar{\mathbf{x}}$ och $(\bar{\mathbf{u}}, \bar{\mathbf{v}})$ det primala respektive det duala problemet.

Följsats 8.4.

Om $\inf \{f(\mathbf{x}) : \mathbf{x} \in X, \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \mathbf{h}(\mathbf{x}) = \mathbf{0}\} = -\infty$ så är $L(\mathbf{u}, \mathbf{v}) = -\infty$ för alla $\mathbf{u} \geq \mathbf{0}$

Följsats 8.5.

Om $\sup \{L(\mathbf{u}, \mathbf{v}) : \mathbf{u} \geq \mathbf{0}\} = \infty$ då har inte det primala problemet någon nåbar lösning.

8.2 Dualitetsgap

Dualitetsgap är något som uppstår när de optimala värdena på målfunktionerna för det primala respektive duala problemet inte är lika. det vill säga ett dualitetsgap är skillnaden mellan det primala och det duala problemets optimala värden. En formell definition följer här:

Definition 8.6. Om $L(\mathbf{u}, \mathbf{v}) < f(\mathbf{x})$ så existerar det ett *dualitetsgap*.

Villkor som garanterar avsaknaden av ett dualitetsgap ges i Sats 9.2. Men innan vi kommer till den satsen behöver vi först några andra resultat.

Lemma 8.7. Farkas Lemma

Låt \mathbf{A} vara en $m \times n$ matris och låt \mathbf{b} vara en n vektor. Då har exakt ett av följande två system en lösning:

System 1: $\mathbf{Ax} = \mathbf{b}$ och $\mathbf{x} \geq \mathbf{0}$ för något $\mathbf{x} \in \mathbf{R}^n$

System 2: $\mathbf{A}^t \mathbf{y} \leq \mathbf{0}$ och $\mathbf{b}^t \mathbf{y} > 0$ för något $\mathbf{y} \in \mathbf{R}^m$

Vi använder oss av Simplexmetoden för fas 1-problem för att bevisa Farkas lemma och därmed skiljer sig vårt bevis från Bazaraas i [5]. Vi hänvisar läsare till den boken för ytterligare ett bevis.

Bevis. Vi gör vårt bevis i två steg.

Steg 1: Visa att (S1) och (S2) inte kan ha lösningar samtidigt.

Antag att System 1 (S1) och System 2 (S2) har lösningar samtidigt. Dvs följande gäller:

$$\mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}$$

$$\mathbf{A}^t \mathbf{y} \leq \mathbf{0}, \mathbf{b}^t \mathbf{y} > 0$$

Om $\mathbf{b}^t \mathbf{y} > 0$ ska gälla så måste \mathbf{b}^t och \mathbf{y} ha samma tecken. Antag att båda är positiva.

$\mathbf{y} > \mathbf{0} \Rightarrow \mathbf{A}^t \leq \mathbf{0}$ för att $\mathbf{A}^t \mathbf{y} \leq \mathbf{0}$ ska gälla och eftersom $\mathbf{y} > \mathbf{0}$ så är $\mathbf{b}^t > \mathbf{0} \Rightarrow \mathbf{b} > \mathbf{0}$. Om $\mathbf{b} > \mathbf{0}$ och $\mathbf{A}^t \leq \mathbf{0} \Rightarrow \mathbf{A} \leq \mathbf{0}$ så får vi att $\mathbf{Ax} \leq \mathbf{0} \neq \mathbf{b}$ ty $\mathbf{b} > \mathbf{0}$, vilket ger en motsägelse från vårt antaganden om att båda systemen hade lösningar. Alltså kan de två systemen inte ha lösningar samtidigt.

Steg 2: Visa att om (S1) saknar lösning så har (S2) lösningar. Vi gör detta genom att studera simplexmetodens fas I-problem för (S1).

Vi vet att optimalvärdet till (P) nedan är större än 0.

$$(P) \quad \text{Minimera} \quad (\mathbf{0}^t, \mathbf{1}^t) \begin{bmatrix} x \\ y \end{bmatrix}$$

$$\text{Då} \quad [A \quad I] \begin{bmatrix} x \\ y \end{bmatrix} = b$$

$$\begin{bmatrix} x \\ y \end{bmatrix} \geq 0$$

(P):s duala problem (D):

Maximera $b^t u$

$$\text{Då} \quad \begin{bmatrix} A^t \\ I^t \end{bmatrix} u \leq \begin{bmatrix} \mathbf{0} \\ \mathbf{1} \end{bmatrix}$$

Eftersom (P) har ändlig positiv lösning så har även det duala problemet (D) ändlig positiv lösning enligt Dualitetssatsen. Dvs det existerar ett u_* så att $A^t u_* \leq \mathbf{0}$ där $b^t u_* > 0$, alltså har (S2) lösning. \square

Lemma 8.8. Gordans lemma

Låt \mathbf{A} vara en $m \times n$ matris. Då har exakt ett av följande två system en lösning

System 1: $\mathbf{Ax} < \mathbf{0}$ för något $\mathbf{x} \in \mathbf{R}^n$

System 2: $\mathbf{A}^t \mathbf{y} = \mathbf{0}$, $\mathbf{y} \geq \mathbf{0}$ för något nollskilt $\mathbf{y} \in \mathbf{R}^m$

Bevis. Vi kan skriva om System 1 på följande ekvivalenta sätt

$$\mathbf{Ax} + \mathbf{e}s \leq \mathbf{0}$$

för något $\mathbf{x} \in \mathbf{R}^n$ och $s > 0$, $s \in \mathbf{R}^k$ och där \mathbf{e} är en vektor av m ettor. Vi kan nu skriva det på samma form som System 1 i Sats 8.7 och vi får då

$$\begin{bmatrix} \mathbf{A} & \mathbf{e} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ s \end{bmatrix} \leq \mathbf{0} \quad \text{och} \quad (0, \dots, 0, 1) \begin{pmatrix} \mathbf{x} \\ s \end{pmatrix} > 0, \quad \text{för något} \quad \begin{pmatrix} \mathbf{x} \\ s \end{pmatrix} \in \mathbf{R}^{n+k}$$

Enligt Lemma 8.7 så ger det associerade System 2 att

$$\begin{bmatrix} \mathbf{A}^t \\ \mathbf{e}^t \end{bmatrix} \mathbf{y} = (0, \dots, 0, 1) \quad \text{och} \quad \mathbf{y} \geq \mathbf{0} \quad \text{för något} \quad \mathbf{y} \in \mathbf{R}^m$$

dvs $\mathbf{A}^t \mathbf{y} = \mathbf{0}$, $\mathbf{e}^t \mathbf{y} = 1$ och $\mathbf{y} \geq \mathbf{0}$ för något $\mathbf{y} \in \mathbf{R}^m$

Detta är ekvivalent med System 2 i Gordans sats och alltså så har exakt ett av systemen lösning. \square

Definition 8.9. Låt $f : S \rightarrow \mathbf{R}^n$ där S är en icke-tom konvex mängd i \mathbf{R}^n . Då är funktionen f konvex på S om

$$f(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha f(x_1) + (1 - \alpha)f(x_2) \quad (8.1)$$

för alla $x_1, x_2 \in \mathbf{R}^n$ och alla tal α där $0 < \alpha < 1$.

Nästa sats behöver vi för att kunna visa senare resultat, eftersom den är en hjälpsats hänvisar vi till [5] för bevis.

Sats 8.10.

Låt S vara en icke-tom, konvex mängd i \mathbf{R}^n och låt $\bar{\mathbf{x}} \in \delta S$ (randen av S). Då existerar det ett hyperplan som stödjer S i $\bar{\mathbf{x}}$; med andra ord det existerar en nollskild vektor \mathbf{p} så att $\mathbf{p}^t(\mathbf{x} - \bar{\mathbf{x}}) \leq 0$ för varje \mathbf{x} som ligger i det slutna höljet av X .

Definition 8.11. En *affin avbildning* (även kallad *affin funktion*) är en sammansättning av en linjär avbildning och en translation (en förflyttning från ett koordinatsystem till ett annat). Geometriskt utgör de affina avbildningarna alla operationer som bevarar räta linjer. Ett grundläggande exempel är:

$$f(x) = ax + b$$

Lemma 8.12.

Låt X vara en icke-tom, konvex mängd i \mathbf{R}^n . Låt $\alpha : \mathbf{R}^n \rightarrow \mathbf{R}^1$ och $g : \mathbf{R}^n \rightarrow \mathbf{R}^m$ vara konvex och låt $\mathbf{h} : \mathbf{R}^n \rightarrow \mathbf{R}^k$ vara en affin funktion; alltså \mathbf{h} är på formen $\mathbf{h}(\mathbf{x}) = \mathbf{A}\mathbf{x} - \mathbf{b}$. Om **System 1** nedan inte har någon lösning \mathbf{x} då har **System 2** en lösning $(u_0, \mathbf{u}, \mathbf{v})$. Det omvända gäller om $u_0 > 0$.

$$\textbf{System 1: } \alpha(\mathbf{x}) < 0, \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \mathbf{h}(\mathbf{x}) = \mathbf{0} \text{ för något } \mathbf{x} \in X$$

$$\textbf{System 2: } u_0\alpha(\mathbf{x}) + \mathbf{u}^t\mathbf{g}(\mathbf{x}) + \mathbf{v}^t\mathbf{h}(\mathbf{x}) \geq 0 \text{ för alla } \mathbf{x} \in X$$

$$(u_0, \mathbf{u}) \geq 0(u_0, \mathbf{u}, \mathbf{v}) \neq \mathbf{0}$$

Bevis. Del 1:

Antag att System 1 inte har någon lösning och studera följande mängd:

$$\Phi = \{(p, \mathbf{q}, \mathbf{r}) : p > \alpha(\mathbf{x}), \mathbf{q} \geq \mathbf{g}(\mathbf{x}), \mathbf{r} = \mathbf{h}(\mathbf{x}) \text{ för något } \mathbf{x} \in X\}$$

Vi noterar att X , α och \mathbf{g} är konvexa och att \mathbf{h} är en affin funktion. Det kan lätt visas att Φ är konvex. Eftersom System 1 inte har någon lösning så är $(0, \mathbf{0}, \mathbf{0}) \notin \Phi$. Enligt Lemma 8.8 och Sats 8.10 så existerar det en nollskild vektor $(u_0, \mathbf{u}, \mathbf{v})$ så att

$$u_0p + \mathbf{u}^t\mathbf{q} + \mathbf{v}^t\mathbf{r} \geq 0 \quad \text{för varje } (p, \mathbf{q}, \mathbf{r}) \text{ i det slutna höljet av } \Phi \quad (8.2)$$

Fixera ett $\mathbf{x} \in X$. Eftersom p och \mathbf{q} kan göras godtyckligt stora så håller (8.2) om $u_0 \geq 0$ och $\mathbf{u} \geq \mathbf{0}$. Dessutom så ligger $(p, \mathbf{q}, \mathbf{r}) = [\alpha(\mathbf{x}), \mathbf{g}(\mathbf{x}), \mathbf{h}(\mathbf{x})]$ i det slutna höljet av Φ . Därför får vi från (8.2) att:

$$u_0\alpha(\mathbf{x}) + \mathbf{u}^t\mathbf{g}(\mathbf{x}) + \mathbf{v}^t\mathbf{h}(\mathbf{x}) \geq 0$$

Eftersom ovanstående olikhet är sann för varje $x \in X$ så har System 2 en lösning.

Del 2:

Vi måste nu visa att om System 2 har en lösning så saknar System 1 lösning. Antag alltså att System 2 har en lösning $(u_0, \mathbf{u}, \mathbf{v})$ så att $u_0 > 0$ och $\mathbf{u} \geq \mathbf{0}$ uppfyller

$$u_0\alpha(\mathbf{x}) + \mathbf{u}^t\mathbf{g}(\mathbf{x}) + \mathbf{v}^t\mathbf{h}(\mathbf{x}) \geq 0 \quad \text{för varje } \mathbf{x} \in X$$

Låt nu $\mathbf{x} \in X$ vara så att $\mathbf{g}(\mathbf{x}) \leq \mathbf{0}$ och $\mathbf{h}(\mathbf{x}) = \mathbf{0}$. Då får vi från ovanstående olikhet, eftersom $\mathbf{u} \geq \mathbf{0}$, att $u_0\alpha(\mathbf{x}) \geq 0$. Då $u_0 > 0, \alpha(\mathbf{x}) \geq 0$ så saknar System 1 lösning. \square

8.3 Stark dualitet

Nästa sats (Starka dualitetssatsen) visar att under lämpliga kovexa antaganden och under ett kvalifikativt villkor så är målfunktionernas optimala värden för det primala och det duala problemet lika. Detta resultat är viktigt då vi fokuserar på att optimera det något enklare duala problemet och på så sätt samtidigt får fram en optimal lösning till det primala problemet.

Sats 8.13. *Starka dualitetssatsen*

Låt X vara en icke-tom konvex mängd i \mathbf{R}^n låt $f : \mathbf{R}^n \rightarrow \mathbf{R}^k$ och $g : \mathbf{R}^n \rightarrow \mathbf{R}^m$ vara konvexa och låt $h : \mathbf{R}^n \rightarrow \mathbf{R}^k$ vara en affin funktion, dvs \mathbf{h} är på formen $\mathbf{h}(\mathbf{x}) = \mathbf{A}\mathbf{x} - \mathbf{b}$. Antag att följande restriktioner håller:

Det existerar ett $\hat{\mathbf{x}} \in X$ så att $\mathbf{g}(\hat{\mathbf{x}}) < \mathbf{0}$ och $\mathbf{h}(\hat{\mathbf{x}}) = \mathbf{0}$ och att $\mathbf{0}$ ligger i det inre av \mathbf{h} där $\mathbf{h}(X) = \{\mathbf{h}(\mathbf{x}) : \mathbf{x} \in X\}$. Då är:

$$\inf\{f(\mathbf{x}) : \mathbf{x} \in X, \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \mathbf{h}(\mathbf{x}) = \mathbf{0}\} = \sup\{L(\mathbf{u}, \mathbf{v}) : \mathbf{u} \geq \mathbf{0}\} \quad (8.3)$$

Dessutom om \inf är ändligt så är $\sup\{L(\mathbf{u}, \mathbf{v}) : \bar{\mathbf{u}} \geq \mathbf{0}\}$ uppnådd i $(\bar{\mathbf{u}}, \bar{\mathbf{v}})$ där $\bar{\mathbf{u}} \geq \mathbf{0}$ och om \inf är uppnådd i $\bar{\mathbf{x}}$ så är $\bar{\mathbf{u}}^t\mathbf{g}(\bar{\mathbf{x}}) = 0$.

Bevis. Låt $\gamma = \inf\{f(\mathbf{x}) : \mathbf{x} \in X, \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \mathbf{h}(\mathbf{x}) = \mathbf{0}\}$. Vi antar att $\gamma < \infty$. Om $\gamma = -\infty$ så får vi från Sats 8.1, Följdsats 8.2, Följdsats 8.3 och Följdsats 8.4 att $\sup\{L(\mathbf{u}, \mathbf{v}) : \mathbf{u} \geq \mathbf{0}\} = -\infty$ och (8.3) sann. Antag därför att γ är ändlig och studera följande system:

$$f(\mathbf{x}) - \gamma < 0 \quad \mathbf{g}(\mathbf{x}) \leq \mathbf{0} \quad \mathbf{h}(\mathbf{x}) = \mathbf{0} \quad \mathbf{x} \in X$$

Enligt definitionen av γ så har inte systemet någon lösning. Vi vet då från Lemma 8.12 att det existerar en nollskild vektor $(u_0, \mathbf{u}, \mathbf{v})$ med $(u_0, \mathbf{u}) \geq \mathbf{0}$ så att:

$$u_0[f(\mathbf{x}) - \gamma] + \mathbf{u}^t\mathbf{g}(\mathbf{x}) + \mathbf{v}^t\mathbf{h}(\mathbf{x}) \geq 0 \quad \text{för alla } \mathbf{x} \in X. \quad (8.4)$$

Vi vill först visa att $u_0 > 0$ genom en motsägelse. Så antag att $u_0 = 0$. Då antar vi att det existerar ett $\hat{\mathbf{x}} \in X$ så att $\mathbf{g}(\hat{\mathbf{x}}) < \mathbf{0}$ och $\mathbf{h}(\hat{\mathbf{x}}) = \mathbf{0}$. Sätt in detta i (8.4), vi får då att $\mathbf{u}^t \mathbf{g}(\hat{\mathbf{x}}) \geq 0$. Men eftersom $\mathbf{g}(\hat{\mathbf{x}}) < \mathbf{0}$ och $\mathbf{u} \geq \mathbf{0}$ så är $\mathbf{u}^t \mathbf{g}(\hat{\mathbf{x}}) \geq 0$ endast möjligt om $\mathbf{u} = \mathbf{0}$. Om $u_0 = 0$ och $\mathbf{u} = \mathbf{0}$ så leder (8.4) till att $\mathbf{v}^t \mathbf{h}(\mathbf{x}) \geq 0$ för alla $\mathbf{x} \in X$. Men då $\mathbf{0}$ ligger i det inre av $\mathbf{h}(X)$ så kan vi ta ett $\mathbf{x} \in X$ så att $\mathbf{h}(\mathbf{x}) = -\lambda \mathbf{v}$, där $\lambda > 0$. Vi får då att $0 \leq \mathbf{v}^t \mathbf{h}(\mathbf{x}) = -\lambda \|\mathbf{v}\|^2$ vilket innebär att $\mathbf{v} = \mathbf{0}$. Vi har nu visat att $u_0 = 0$ innebär att $(u_0, \mathbf{u}, \mathbf{v}) = \mathbf{0}$ vilket inte är möjligt. Alltså måste $u_0 > 0$. Om vi delar (8.4) med u_0 och sätter $\frac{\mathbf{u}}{u_0} = \bar{\mathbf{u}}$ och $\frac{\mathbf{v}}{u_0} = \bar{\mathbf{v}}$ så får vi:

$$f(\mathbf{x}) + \bar{\mathbf{u}}^t \mathbf{g}(\mathbf{x}) + \bar{\mathbf{v}}^t \mathbf{h}(\mathbf{x}) \geq \gamma \quad \text{för alla } \mathbf{x} \in X \quad (8.5)$$

Detta visar att $L(\bar{\mathbf{u}}, \bar{\mathbf{v}}) = \inf\{f(\mathbf{x}) + \bar{\mathbf{u}}^t \mathbf{g}(\mathbf{x}) + \bar{\mathbf{v}}^t \mathbf{h}(\mathbf{x}) : \mathbf{x} \in X\} \geq \gamma$. Vi får då från Sats 8.1 att $L(\bar{\mathbf{u}}, \bar{\mathbf{v}}) = \gamma$ och att $(\bar{\mathbf{u}}, \bar{\mathbf{v}})$ löser det duala problemet. Det sista steget i beviset är att visa att $\bar{\mathbf{u}}^t \mathbf{g}(\bar{\mathbf{x}}) = 0$. Antag att $\bar{\mathbf{x}}$ är en lösning till det primala problemet, det vill säga $\bar{\mathbf{x}} \in X$, $\mathbf{g}(\bar{\mathbf{x}}) \leq \mathbf{0}$, $\mathbf{h}(\bar{\mathbf{x}}) = \mathbf{0}$ och att $f(\bar{\mathbf{x}}) = \gamma$. Från (8.5) får vi, om vi låter $\mathbf{x} = \bar{\mathbf{x}}$, att $\bar{\mathbf{u}}^t \mathbf{g}(\bar{\mathbf{x}}) \geq 0$. Men eftersom $\bar{\mathbf{u}} \geq \mathbf{0}$ och $\mathbf{g}(\bar{\mathbf{x}}) \leq \mathbf{0}$ så måste $\bar{\mathbf{u}}^t \mathbf{g}(\bar{\mathbf{x}}) = 0$. \square

Det är inte alltid möjligt att finna en explicit lösning till alla problem, exemplet nedan visar att vi i sådana fall ändå kan skaffa oss en idé till hur en lösning kan se ut. Vi kan med nedanstående metod stänga in lösningen inom ett intervall och därmed uppskatta en ungefärlig lösning.

Exempel 6.

$$\begin{aligned} \min f(x_1, x_2) &= 2x_1^2 + 5x_2^2 - 6x_1 \\ \text{Då} \quad \quad \quad &3x_1^2 - 2x_2 \leq 6 \end{aligned}$$

Bilda

$$\begin{aligned} L(u) &= \min\{2x_1^2 + 5x_2^2 - 6x_1 + u(3x_1^2 - 2x_2 - 6)\} = \\ &= \min\{2x_1^2 + 3ux_1^2 - 6x_1\} + \min\{5x_2^2 - 2ux_2 - 6u\} = \\ &= (3u + 2)\left(\frac{3}{2 + 3u}\right)^2 - 6\frac{3}{2 + 3u} + 5\left(\frac{u}{5}\right)^2 - 2u\left(\frac{u}{5}\right) - 6u = \\ &= -\frac{9}{2 + 3u} - \frac{u^2}{5} - 6u \end{aligned}$$

Vi ser att

$$L(0) = -\frac{9}{2}$$

Så att

$$\begin{aligned} x_1 &= \frac{3}{2}, \quad x_2 = 0 \Rightarrow \\ \Rightarrow 3x_1^2 - 2x_2 &= 6\frac{3}{4} > 6 \end{aligned}$$

Så istället använder vi

$$L(1) = -\frac{9}{5} - \frac{1}{5} - 6 = -8$$

Dvs

$$x_1 = \frac{3}{5} \quad x_2 = \frac{1}{5}$$

Och en snabbkontroll ger att

$$3x_1^2 - 2x_2 = 3\left(\frac{3}{5}\right)^2 - 2\frac{1}{5} = 1 \leq 6$$

Värdet i f är alltså:

$$f\left(\frac{3}{5}, \frac{1}{5}\right) = 2\frac{9}{25} + 5\frac{1}{25} - 6\frac{3}{5} = -\frac{67}{25}$$

Vi kan alltså sluta oss till att

$$-\frac{9}{2} \leq f(x^*) \leq -\frac{67}{25}$$

△

9 Sadelpunktskriterier och KKT-villkoren

9.1 Sadelpunktskriterier

Starka dualitetssatsen i föregående avsnitt visar på att under konvexitetsantaganden och under en lämplig restriktion kommer värdena för det primala och det duala problemets målfunktioner att vara lika i optimum. Faktiskt så är existensen av en sadelpunkt ett nödvändigt och tillräckligt villkor för att föregående sats egenskaper ska hålla. Givet det primala problemet P definierar vi *Lagrangefunktionen*

$$\Lambda(\mathbf{x}, \mathbf{u}, \mathbf{v}) = f(\mathbf{x}) + \mathbf{u}^t \mathbf{g}(\mathbf{x}) + \mathbf{v}^t \mathbf{h}(\mathbf{x})$$

Definition 9.1. En lösning $(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}})$ till Lagrangefunktionen kallas för *sadelpunkt* om $\bar{\mathbf{x}} \in X$, $\bar{\mathbf{u}} \geq \mathbf{0}$ och

$$\Lambda(\bar{\mathbf{x}}, \mathbf{u}, \mathbf{v}) \leq \Lambda(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}}) \leq \Lambda(\mathbf{x}, \bar{\mathbf{u}}, \bar{\mathbf{v}}) \text{ för alla } \mathbf{x} \in X \text{ och alla } (\mathbf{u}, \mathbf{v}) \text{ där } \mathbf{u} \geq \mathbf{0} \quad (9.1)$$

Vi har därför att $\bar{\mathbf{x}}$ minimerar Lagrangefunktionen över X när (\mathbf{u}, \mathbf{v}) är fix vid $(\bar{\mathbf{u}}, \bar{\mathbf{v}})$ och att $(\bar{\mathbf{u}}, \bar{\mathbf{v}})$ maximerar Lagrangefunktionen över alla (\mathbf{u}, \mathbf{v}) där $\mathbf{u} \geq \mathbf{0}$ när \mathbf{x} är fix vid $\bar{\mathbf{x}}$. Satsen nedan karakteriserar en sadelpunktslösning och visar att dess existens är ett nödvändigt villkor för att det inte ska finnas något dualitetsgap.

Sats 9.2. Sadelpunktsoptimalitet och avsaknaden av dualitetsgap

En lösning $(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}})$ där $\bar{\mathbf{x}} \in X$ och $\bar{\mathbf{u}} \geq \mathbf{0}$ är en sadelpunkt för Lagrangefunktionen $\Lambda(\mathbf{x}, \mathbf{u}, \mathbf{v}) = f(\mathbf{x}) + \mathbf{u}^t \mathbf{g}(\mathbf{x}) + \mathbf{v}^t \mathbf{h}(\mathbf{x})$ om och endast om

(i) $\Lambda(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}}) = \min\{\Lambda(\mathbf{x}, \bar{\mathbf{u}}, \bar{\mathbf{v}}) : \mathbf{x} \in X\}$

(ii) $\mathbf{g}(\bar{\mathbf{x}}) \leq \mathbf{0}$, $\mathbf{h}(\bar{\mathbf{x}}) = \mathbf{0}$ och

(iii) $\bar{\mathbf{u}}^t \mathbf{g}(\bar{\mathbf{x}}) = 0$

Dessutom är $(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}})$ en sadelpunkt för Lagrangefunktionen om och endast om $\bar{\mathbf{x}}$ är en optimal lösning till det primala problemet P och $(\bar{\mathbf{u}}, \bar{\mathbf{v}})$ är en optimal lösning till det duala problemet D utan att det existerar ett dualitetsgap, det vill säga $f(\bar{\mathbf{x}}) = L(\bar{\mathbf{u}}, \bar{\mathbf{v}})$.

Bevis. Antag att $(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}})$ är en sadelpunkt till Lagrangefunktionen Λ . Enligt definitionen så måste villkor (i) vara sant. Dessutom har vi från (9.1) att

$$f(\bar{\mathbf{x}}) + \bar{\mathbf{u}}^t \mathbf{g}(\bar{\mathbf{x}}) + \bar{\mathbf{v}}^t \mathbf{h}(\bar{\mathbf{x}}) \geq f(\bar{\mathbf{x}}) + \mathbf{u}^t \mathbf{g}(\bar{\mathbf{x}}) + \mathbf{v}^t \mathbf{h}(\bar{\mathbf{x}}) \text{ för alla } (\mathbf{u}, \mathbf{v}) \text{ med } \mathbf{u} \geq \mathbf{0} \quad (9.2)$$

Detta innebär att $\mathbf{g}(\bar{\mathbf{x}}) \leq \mathbf{0}$ och $\mathbf{h}(\bar{\mathbf{x}}) = \mathbf{0}$ måste gälla för om man väljer en lämplig komponent av \mathbf{u} eller \mathbf{v} av oändlig storlek så håller inte (9.2). Om vi nu sätter $\mathbf{u} = \mathbf{0}$ i (9.2) så får vi att $\bar{\mathbf{u}}^t \mathbf{g}(\bar{\mathbf{x}}) \geq 0$. Notera att $\bar{\mathbf{u}} \geq \mathbf{0}$ och $\mathbf{g}(\bar{\mathbf{x}}) \leq \mathbf{0}$ ger att $\bar{\mathbf{u}}^t \mathbf{g}(\bar{\mathbf{x}}) \leq 0$ alltså måste $\bar{\mathbf{u}}^t \mathbf{g}(\bar{\mathbf{x}}) = 0$. Därför så håller villkoren (i), (ii) och (iii).

Antag nu motsatsen, det vill säga att $(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}})$ där $\bar{\mathbf{x}} \in X$ och $\bar{\mathbf{u}} \geq \mathbf{0}$ så att villkoren (i), (ii), (iii) är givna. Då gäller $\Lambda(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}}) \leq \Lambda(\mathbf{x}, \bar{\mathbf{u}}, \bar{\mathbf{v}})$ för alla $\bar{\mathbf{x}} \in X$ med egenskap (i). Dessutom så är $\Lambda(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}}) = f(\bar{\mathbf{x}}) + \bar{\mathbf{u}}^t \mathbf{g}(\bar{\mathbf{x}}) + \bar{\mathbf{v}}^t \mathbf{h}(\bar{\mathbf{x}}) = f(\bar{\mathbf{x}}) \geq f(\bar{\mathbf{x}}) + \mathbf{u}^t \mathbf{g}(\bar{\mathbf{x}}) + \mathbf{v}^t \mathbf{h}(\bar{\mathbf{x}}) = \Lambda(\bar{\mathbf{x}}, \mathbf{u}, \mathbf{v})$ för alla $\mathbf{u} \geq \mathbf{0}$, eftersom $\mathbf{g}(\bar{\mathbf{x}}) \leq \mathbf{0}$ och $\mathbf{h}(\bar{\mathbf{x}}) = \mathbf{0}$. Därför är $(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}})$ en sadelpunkt.

Nästa steg, antag igen att $(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}})$ är en sadelpunkt. Enligt (ii) så är $\bar{\mathbf{x}}$ nåbar lösning till problem P. Eftersom $\bar{\mathbf{u}} \geq \mathbf{0}$ har vi också att $(\bar{\mathbf{u}}, \bar{\mathbf{v}})$ är en nåbar lösning till det duala problemet D. Dessutom enligt villkor (i), (ii) och (iii) så är $L(\bar{\mathbf{u}}, \bar{\mathbf{v}}) = \Lambda(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}}) = f(\bar{\mathbf{x}}) + \bar{\mathbf{u}}^t \mathbf{g}(\bar{\mathbf{x}}) + \bar{\mathbf{v}}^t \mathbf{h}(\bar{\mathbf{x}}) = f(\bar{\mathbf{x}})$. Från Sats 8.1, Följdsats 8.2 och Följdsats 8.3 så ser vi att $\bar{\mathbf{x}}$ och $(\bar{\mathbf{u}}, \bar{\mathbf{v}})$ löser problem P respektive problem D utan att det existerar något dualitetsgap.

Avslutningsvis så antar vi att $\bar{\mathbf{x}}$ och $(\bar{\mathbf{u}}, \bar{\mathbf{v}})$ är optimala lösningar till problem P respektive problem D, där $f(\bar{\mathbf{x}}) = L(\bar{\mathbf{u}}, \bar{\mathbf{v}})$. Därför har vi att $\bar{\mathbf{x}} \in X$, $\mathbf{g}(\bar{\mathbf{x}}) \leq \mathbf{0}$, $\mathbf{h}(\bar{\mathbf{x}}) = \mathbf{0}$ och $\bar{\mathbf{u}} \geq \mathbf{0}$. Dessutom vet vi sedan tidigare resultat att

$$\begin{aligned} L(\bar{\mathbf{u}}, \bar{\mathbf{v}}) &= \min\{f(\mathbf{x}) + \bar{\mathbf{u}}^t \mathbf{g}(\mathbf{x}) + \bar{\mathbf{v}}^t \mathbf{h}(\mathbf{x}) : \mathbf{x} \in X\} \\ &\leq f(\bar{\mathbf{x}}) + \bar{\mathbf{u}}^t \mathbf{g}(\bar{\mathbf{x}}) + \bar{\mathbf{v}}^t \mathbf{h}(\bar{\mathbf{x}}) = f(\bar{\mathbf{x}}) + \bar{\mathbf{u}}^t \mathbf{g}(\bar{\mathbf{x}}) \leq f(\bar{\mathbf{x}}) \end{aligned}$$

Men vi har att $L(\bar{\mathbf{u}}, \bar{\mathbf{v}}) = f(\bar{\mathbf{x}})$ enligt antagandet. Alltså så måste likheter gälla ovan. Särskilt om $\bar{\mathbf{u}}^t \mathbf{g}(\bar{\mathbf{x}}) = 0$ för då $\Lambda(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}}) = f(\bar{\mathbf{x}}) = L(\bar{\mathbf{u}}, \bar{\mathbf{v}}) = \min\{\Lambda(\mathbf{x}, \bar{\mathbf{u}}, \bar{\mathbf{v}}) : \mathbf{x} \in X\}$. Därför håller villkor (i), (ii) och (iii) om $\bar{\mathbf{x}} \in X$ och $\bar{\mathbf{u}} \geq \mathbf{0}$ och alltså så är $(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}})$ en sadelpunkt. \square

9.2 Karush-Kuhn-Tuckers villkor

Karush-Kuhn-Tuckers villkor, eller KKT-villkoren som vi kommer att förkorta det till, är villkor som måste vara uppfyllda för att en punkt ska vara en optimallösning till ett optimeringsproblem. Villkoren är nödvändiga men inte tillräckliga, dvs om villkoren är uppfyllda behöver inte punkten vara optimum, men optimum uppfyller villkoren. Vi introducerar i det här avsnittet KKT-villkoren för att i nästa avsnitt kunna visa relationen mellan dessa och sadelpunktsoptimalitet, vi kommer att visa att sadelpunktsoptimalitet medför att KKT-villkoren är uppfyllda och vice versa.

Definition 9.3. *Gradienten*, eller gradientvektorn $\nabla f(\mathbf{x})$ definieras som:

$$\nabla f(\mathbf{x}) = \left(\frac{\partial f(\mathbf{x})}{\partial x_1}, \dots, \frac{\partial f(\mathbf{x})}{\partial x_n} \right)^t$$

Exempel 7. Använd Lagrangedualitet för att lösa

$$\begin{aligned} \min \quad & f(x_1, x_2) = x_1^2 + 3x_2^2 \\ \text{då} \quad & x_1 + 2x_2 \geq 1 \implies -x_1 - 2x_2 + 1 \leq 0 \end{aligned}$$

Lagranges duala funktion blir då $L(u) = \min_x \{x_1^2 + 3x_2^2 + u(-x_1 - 2x_2 + 1)\}$

$$\begin{aligned} \nabla_x L(x_1, x_2) &= (2x_1 - u, 6x_2 - 2u) = 0 \\ \implies \quad \bar{x}_1(u) &= \frac{u}{2} \quad \bar{x}_2(u) = \frac{u}{3} \\ \implies L(u) &= \frac{u^2}{4} + \frac{3u^2}{9} - u \frac{u}{2} - 2u \frac{u}{3} + u = -\frac{7}{12}u^2 + u \end{aligned}$$

För att hitta $\max_{u \geq 0} L(u)$ så deriverar vi:

$$L'(u) = -\frac{7}{6}u + 1 = 0 \implies \bar{u} = \frac{6}{7}$$

$$L\left(\frac{6}{7}\right) = \frac{3}{7} \implies$$

$$x^* = \bar{x}(\bar{u}) = \left(\frac{1}{2} \cdot \frac{3}{7}, \frac{1}{3} \cdot \frac{6}{7} \right) = \left(\frac{3}{7}, \frac{2}{7} \right) \implies$$

$$f(x^*) = \left(\frac{3}{7} \right)^2 + 3 \left(\frac{2}{7} \right)^2 = \frac{3}{7}$$

och vi ser här att $f(x^*) = L(\bar{u})$ vilket Sats 8.13 säger. \triangle

Sats 9.4. KKT-villkoren

Låt X vara en icke-tom öppen mängd i \mathbf{R}^n och låt $f : \mathbf{R}^n \rightarrow \mathbf{R}^k$ och $g_i : \mathbf{R}^n \rightarrow \mathbf{R}^k$ för $i = 1, \dots, m$. Betrakta problem P som att minimera $f(\mathbf{x})$ under villkoren $\mathbf{x} \in X$ och $g_i(\mathbf{x}) \leq 0$ för $i = 1, \dots, m$. Låt $\bar{\mathbf{x}}$ vara en möjlig lösning och sätt $I = \{i ; g_i(\bar{\mathbf{x}}) = 0\}$. Antag att f och g_i , $i \in I$, är differentierbara i $\bar{\mathbf{x}}$ och att g_i är kontinuerlig i $\bar{\mathbf{x}}$ då $i \notin I$. Anta även att $\nabla g_i(\bar{\mathbf{x}})$ är linjärt oberoende för $i \in I$. Om $\bar{\mathbf{x}}$ är en lokal lösning till problem P så existerar det en skalär u_i för $i \in I$ så att

$$\nabla f(\bar{\mathbf{x}}) + \sum_{i \in I} u_i \nabla g_i(\bar{\mathbf{x}}) = 0$$

$$u_i \geq 0 \quad i \in I$$

Tillsammans med ovanstående antaganden så får vi att g_i för varje $i \in I$ också är differentierbar i $\bar{\mathbf{x}}$ och vi kan skriva föregående villkor på följande ekvivalenta sätt:

$$\begin{aligned} \nabla f(\bar{\mathbf{x}}) + \sum_{i=1}^m u_i \nabla g_i(\bar{\mathbf{x}}) &= \mathbf{0} \\ u_i g_i(\bar{\mathbf{x}}) &= 0 \quad (i = 1, \dots, m) \\ u_i &\geq 0 \quad (i = 1, \dots, m) \end{aligned}$$

Bevis. Det existerar skalärer u_0 och \hat{u}_i för $i \in I$, inte alla lika med 0, så att

$$\begin{aligned} u_0 \nabla f(\bar{\mathbf{x}}) + \sum_{i \in I} \hat{u}_i \nabla g_i(\bar{\mathbf{x}}) &= \mathbf{0} \\ u_0, \hat{u}_i &\geq 0 \quad i \in I \end{aligned} \tag{9.3}$$

Vi observerar här att u_0 måste vara större än 0 för om $u_0 = 0$ så skulle (9.3) motsäga vårt antagande om att $\nabla g_i(\bar{\mathbf{x}})$ är linjärt oberoende för $i \in I$. Första delen av satsen följer från att vi låter $u_i = \frac{\hat{u}_i}{u_0}$ för varje $i \in I$. Den ekvivalenta omskrivningen av de nödvändiga villkoren följer från att vi låter $u_i = 0$ för de $i \notin I$.

□

9.3 Relationen mellan sadelpunktsoptimalitet och KKT-villkoren

Satsen nedan visar att om KKT-villkoren (Sats 9.4) är uppfyllda så medför det sadelpunktsoptimalitet (Sats 9.2). Satsen visar också att det omvända gäller.

Sats 9.5.

Låt $S = \{\mathbf{x} \in X; \mathbf{g}(\bar{\mathbf{x}}) \leq 0, \mathbf{h}(\bar{\mathbf{x}}) = \mathbf{0}\}$ där X är en icke-tom konvex mängd i \mathbf{R}^n och betrakta problem P som att minimera $f(\mathbf{x})$ under villkoret $\mathbf{x} \in S$. Antag att $\bar{\mathbf{x}} \in S$ uppfyller KKT-villkoren, med andra ord $\exists \bar{\mathbf{u}} \geq \mathbf{0}$ och $\bar{\mathbf{v}}$ så att

$$\begin{aligned} \nabla f(\bar{\mathbf{x}}) + \nabla \mathbf{g}(\bar{\mathbf{x}})^t \bar{\mathbf{u}} + \nabla \mathbf{h}(\bar{\mathbf{x}})^t \bar{\mathbf{v}} &= \mathbf{0} \\ \bar{\mathbf{u}}^t \mathbf{g}(\bar{\mathbf{x}}) &= 0 \end{aligned} \tag{9.4}$$

Låt $I = \{i; g_i(\bar{\mathbf{x}}) = 0\}$ och antag att f , g_i då $i \in I$ är konvexa i $\bar{\mathbf{x}}$. Antag också att $\bar{v}_i \neq 0$, då är h_i affin. Då är $(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}})$ en sadelpunkt för Lagrange-funktionen

$$\Lambda(\mathbf{x}, \mathbf{u}, \mathbf{v}) = f(\mathbf{x}) + \mathbf{u}^t \mathbf{g}(\mathbf{x}) + \mathbf{v}^t \mathbf{h}(\mathbf{x})$$

Om vi omvänt antar att $(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}})$ där $\bar{\mathbf{x}}$ ligger i det inre av X och $\bar{\mathbf{u}} \geq 0$ är en sadelpunktslösning så är $\bar{\mathbf{x}}$ en möjlig lösning till problem P och $(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}})$ uppfyller KKT villkoren i (9.4).

Bevis. Antag att $(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}})$ där $\bar{\mathbf{x}} \in S$ och $\bar{\mathbf{u}} \geq 0$ uppfyller KKT i (9.4). Genom antagandet att f och g_i är konvexa i X då $i \in I$ och att h_i är affin då $\bar{v}_i \neq 0$ så får vi att

$$f(\mathbf{x}) \geq f(\bar{\mathbf{x}}) + \nabla f(\bar{\mathbf{x}})^t(\mathbf{x} - \bar{\mathbf{x}}) \quad (9.5)$$

$$g_i(\mathbf{x}) \geq g_i(\bar{\mathbf{x}}) + \nabla g_i(\bar{\mathbf{x}})^t(\mathbf{x} - \bar{\mathbf{x}}) \quad \text{för } i \in I \quad (9.6)$$

$$h_i(\mathbf{x}) \geq h_i(\bar{\mathbf{x}}) + \nabla h_i(\bar{\mathbf{x}})^t(\mathbf{x} - \bar{\mathbf{x}}) \quad \text{för } i = 1, 2, \dots, k \text{ och } \bar{v}_i \neq 0 \quad (9.7)$$

för alla $\mathbf{x} \in X$.

Om vi multiplicerar (9.6) med $\bar{u}_i \geq 0$ och (9.7) med \bar{v}_i så får vi

$$g_i(\mathbf{x})\bar{u}_i \geq g_i(\bar{\mathbf{x}})\bar{u}_i + \nabla g_i(\bar{\mathbf{x}})^t(\mathbf{x} - \bar{\mathbf{x}})\bar{u}_i \quad (9.8)$$

$$h_i(\mathbf{x})\bar{v}_i \geq h_i(\bar{\mathbf{x}})\bar{v}_i + \nabla h_i(\bar{\mathbf{x}})^t(\mathbf{x} - \bar{\mathbf{x}})\bar{v}_i \quad (9.9)$$

Addera (9.8) och (9.9) med (9.5) och använd (9.4), vi får då

$$\begin{aligned} f(\mathbf{x}) + g_i(\mathbf{x})\bar{u}_i + h_i(\mathbf{x})\bar{v}_i &\geq f(\bar{\mathbf{x}}) + g_i(\bar{\mathbf{x}})\bar{u}_i + h_i(\bar{\mathbf{x}})\bar{v}_i \\ \implies \Lambda(\mathbf{x}, \bar{\mathbf{u}}, \bar{\mathbf{v}}) &\geq \Lambda(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}}) \quad \text{för alla } \mathbf{x} \in X \end{aligned}$$

Eftersom $\mathbf{g}(\bar{\mathbf{x}}) \leq \mathbf{0}$, $\mathbf{h}(\bar{\mathbf{x}}) = \mathbf{0}$ och $\bar{\mathbf{u}}^t \mathbf{g}(\bar{\mathbf{x}}) = 0$ så följer att $\Lambda(\bar{\mathbf{x}}, \mathbf{u}, \mathbf{v}) = f(\bar{\mathbf{x}}) + \mathbf{u}^t \mathbf{g}(\bar{\mathbf{x}}) + \mathbf{v}^t \mathbf{h}(\bar{\mathbf{x}}) \leq f(\bar{\mathbf{x}}) = \Lambda(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}})$ för alla (\mathbf{u}, \mathbf{v}) med $\mathbf{u} \geq \mathbf{0}$ och vi ser att $\Lambda(\bar{\mathbf{x}}, \mathbf{u}, \mathbf{v}) \leq \Lambda(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}}) \leq \Lambda(\mathbf{x}, \bar{\mathbf{u}}, \bar{\mathbf{v}})$ vilket innebär att $\Lambda(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}})$ uppfyller villkoren för en sadelpunkt.

För att bevisa det omvända antar vi att $(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}})$ är en sadelpunkt där $\bar{\mathbf{x}}$ ligger i det inre av X och att $\bar{\mathbf{u}} \geq \mathbf{0}$. Eftersom $\Lambda(\bar{\mathbf{x}}, \mathbf{u}, \mathbf{v}) \leq \Lambda(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}})$ för alla $\mathbf{u} \geq \mathbf{0}$ och alla \mathbf{v} så har vi från (9.2) i Sats 9.2 att $\mathbf{g}(\bar{\mathbf{x}}) \leq \mathbf{0}$, $\mathbf{h}(\bar{\mathbf{x}}) = \mathbf{0}$ och

$\bar{\mathbf{u}}^t \mathbf{g}(\bar{\mathbf{x}}) = \mathbf{0}$. Detta visar att $\bar{\mathbf{x}}$ är en möjlig lösning till problem P. Eftersom $\Lambda(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}}) \leq \Lambda(\mathbf{x}, \bar{\mathbf{u}}, \bar{\mathbf{v}})$ för alla $\mathbf{x} \in X$ så löser $\bar{\mathbf{x}}$ problemet att minimera $\Lambda(\mathbf{x}, \bar{\mathbf{u}}, \bar{\mathbf{v}})$ under villkoret $\mathbf{x} \in X$ och eftersom $\bar{\mathbf{x}}$ ligger i det inre av X så gäller att $\nabla_x \Lambda(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}}) = \mathbf{0}$, det vill säga $\nabla f(\bar{\mathbf{x}}) + \nabla \mathbf{g}(\bar{\mathbf{x}})^t \bar{\mathbf{u}} + \nabla \mathbf{h}(\bar{\mathbf{x}})^t \bar{\mathbf{v}} = \mathbf{0}$ och därför håller (9.4).

□

10 Duala funktionens egenskaper

Vi har tidigare diskuterat relationen mellan de primala och de duala problemen. Vi har också visat att optimallösningarna ger samma värden på den primala och den duala målfunktionen. För att kunna få fram en lösning på det duala problemet måste vi känna till några egenskaper som den duala funktionen har. Vi kommer från och med nu anta att mängden X är kompakt, vi kommer även att, för att göra notationen enklare och mer lättläst, låta vektorn \mathbf{w} vara en kombination av vektorerna \mathbf{u} och \mathbf{v} och vi låter β vara en kombination av funktionerna \mathbf{g} och \mathbf{h} . I vår nästa sats så kommer vi att visa att L är konkav.

Sats 10.1.

Låt X vara en icke-tom, kompakt mängd i \mathbf{R}^n och låt $f : \mathbf{R}^n \rightarrow \mathbf{R}^m$ och $\beta : \mathbf{R}^n \rightarrow \mathbf{R}^{m+k}$ vara kontinuerliga. Om L är definierad som

$$\mathbf{L}(\mathbf{w}) = \inf \{f(\mathbf{x}) + \mathbf{w}^t \beta; \mathbf{x} \in X\}$$

så är L konkav på \mathbf{R}^{m+k} .

Bevis. Eftersom \mathbf{f} och β är kontinuerliga och X är kompakt så är L ändlig överallt på \mathbf{R}^{m+k} . Låt $\mathbf{w}_1, \mathbf{w}_2 \in \mathbf{R}^{m+k}$ och låt $\lambda \in (0, 1)$. Vi har då att

$$\begin{aligned} L[\lambda \mathbf{w}_1 + (1 - \lambda) \mathbf{w}_2] &= \inf \{f(\mathbf{x}) + [\lambda \mathbf{w}_1 + (1 - \lambda) \mathbf{w}_2]^t \beta; \mathbf{x} \in X\} \\ &= \inf \{ \lambda [f(\mathbf{x}) + \mathbf{w}_1^t \beta(\mathbf{x})] + (1 - \lambda) [f(\mathbf{x}) + \mathbf{w}_2^t \beta(\mathbf{x})]; \mathbf{x} \in X \} \\ &\geq \lambda \inf \{f(\mathbf{x}) + \mathbf{w}_1^t \beta(\mathbf{x}); \mathbf{x} \in X\} + (1 - \lambda) \inf \{f(\mathbf{x}) + \mathbf{w}_2^t \beta(\mathbf{x}); \mathbf{x} \in X\} \\ &= \lambda L(\mathbf{w}_1) + (1 - \lambda) L(\mathbf{w}_2) \end{aligned}$$

alltså så är L konkav. □

Eftersom L är konkav så är ett lokalt optimum också ett globalt optimum, vilket är bra eftersom vi vill maximera L . Att lösa det duala problemet är dock inte helt trivialt, den största svårigheten ligger i att den duala funktionen inte alltid är möjlig att nå ty vi kan bara uppskatta L i en punkt efter att vi har löst ett minimeringsdelproblem. För att göra detta så studerar vi hur differentierbarheten och subdifferentierbarheten för L ser ut. Optimeringsproblemet att uppskatta L refereras ibland som *Lagranges duala delproblem*.

L har vi definierat som $L(\mathbf{w}) = \{f(\mathbf{x}) + \mathbf{w}^t \beta(\mathbf{x}); \mathbf{x} \in X\}$ där X är en kompakt mängd i \mathbf{R}^n , vi inför nu en mängd $X(\mathbf{w})$ av optimala lösningar till Lagranges duala delproblem så att

$$X(\mathbf{w}) = \{\mathbf{y}; \mathbf{y} \text{ minimerar } f(\mathbf{x}) + \mathbf{w}^t \beta(\mathbf{x}) \text{ då } \mathbf{x} \in X\}$$

Differentierbarheten hos L i någon given punkt $\bar{\mathbf{w}}$ beror på $X(\bar{\mathbf{w}})$:s element. Om $X(\bar{\mathbf{w}})$ är en singleton, dvs den består av bara ett element, så kommer Sats 10.5 nedan visa att L är differentierbar i $\bar{\mathbf{w}}$, men först behöver vi några definitioner och ett lemma.

Definition 10.2. Låt S vara en icke-tom konvex mängd i \mathbf{R}^n och låt $f : S \rightarrow \mathbf{R}^m$ vara konvex. Då kallas ξ *subgradient till f* i punkten $\bar{\mathbf{x}} \in S$ om

$$f(\mathbf{x}) \geq f(\bar{\mathbf{x}}) + \xi^t(\mathbf{x} - \bar{\mathbf{x}}) \quad \text{för alla } \mathbf{x} \in S$$

Om f istället är konkav så är ξ en *subgradient till f* i $\bar{\mathbf{x}} \in S$ om

$$f(\mathbf{x}) \leq f(\bar{\mathbf{x}}) + \xi^t(\mathbf{x} - \bar{\mathbf{x}}) \quad \text{för alla } \mathbf{x} \in S$$

Definition 10.3. Mängden av alla subgradients till f i punkten $\bar{\mathbf{x}}$ kallas för *subdifferential* av f i $\bar{\mathbf{x}}$.

Lemma 10.4.

Låt X vara en icke-tom, kompakt mängd i \mathbf{R}^n och låt $f : \mathbf{R}^n \rightarrow \mathbf{R}^m$ och $\beta : \mathbf{R}^n \rightarrow \mathbf{R}^{m+k}$ vara kontinuerliga. Låt $\bar{\mathbf{w}} \in \mathbf{R}^{m+k}$ och antag att $X(\bar{\mathbf{w}}) = \{\bar{\mathbf{x}}\}$, dvs en singleton. Antag också att $\mathbf{w}_j \rightarrow \bar{\mathbf{w}}$ och låt $\mathbf{x}_j \in X(\mathbf{w}_j)$ för varje j . Då kommer $\mathbf{x}_j \rightarrow \bar{\mathbf{x}}$

Sats 10.5. Låt X vara en icke-tom, kompakt mängd i \mathbf{R}^n och låt $f : \mathbf{R}^n \rightarrow \mathbf{R}^m$ och $\beta : \mathbf{R}^n \rightarrow \mathbf{R}^{m+k}$ vara kontinuerliga. Låt $\bar{\mathbf{w}} \in \mathbf{R}^{m+k}$ och antag att $X(\bar{\mathbf{w}}) = \{\bar{\mathbf{x}}\}$. Då är L differentierbar i $\bar{\mathbf{w}}$ och gradienten $\nabla L(\bar{\mathbf{w}}) = \beta(\bar{\mathbf{x}})$.

Bevis. Eftersom f , β kontinuerliga och X kompakt så existerar det för något givet \mathbf{w} ett $\mathbf{x}_w \in X(\mathbf{w})$. Från definitionen av L så vet vi att följande gäller

$$L(\mathbf{w}) - L(\bar{\mathbf{w}}) \leq f(\bar{\mathbf{x}}) + \mathbf{w}^t \beta(\bar{\mathbf{x}}) - f(\bar{\mathbf{x}}) - \bar{\mathbf{w}}^t \beta(\bar{\mathbf{x}}) = (\mathbf{w} - \bar{\mathbf{w}})^t \beta(\bar{\mathbf{x}}) \quad (10.1)$$

och

$$L(\bar{\mathbf{w}}) - L(\mathbf{w}) \leq f(\mathbf{x}_w) + \bar{\mathbf{w}}^t \beta(\mathbf{x}_w) - f(\mathbf{x}_w) - \mathbf{w}^t \beta(\mathbf{x}_w) = (\mathbf{w} - \bar{\mathbf{w}})^t \beta(\mathbf{x}_w) \quad (10.2)$$

(10.1) och (10.2) \implies

$$\begin{aligned} 0 &\geq L(\mathbf{w}) - L(\bar{\mathbf{w}}) - (\mathbf{w} - \bar{\mathbf{w}})^t \beta(\bar{\mathbf{x}}) \geq (\mathbf{w} - \bar{\mathbf{w}})^t (\beta(\bar{\mathbf{x}}) - \beta(\mathbf{x}_w)) \geq \\ &\geq - \|\mathbf{w} - \bar{\mathbf{w}}\| \|\beta(\mathbf{x}_w) - \beta(\bar{\mathbf{x}})\| \implies \\ - \|\beta(\mathbf{x}_w) - \beta(\bar{\mathbf{x}})\| &\leq \frac{L(\mathbf{w}) - L(\bar{\mathbf{w}}) - (\mathbf{w} - \bar{\mathbf{w}})^t \beta(\bar{\mathbf{x}})}{\|\mathbf{w} - \bar{\mathbf{w}}\|} \leq 0 \end{aligned} \quad (10.3)$$

När $\mathbf{w} \rightarrow \bar{\mathbf{w}}$ så $\mathbf{x}_w \rightarrow \bar{\mathbf{x}}$ enligt Lemma 10.4 och eftersom β är kontinuerlig så (se [7]) $\beta(\mathbf{x}_w) \rightarrow \beta(\bar{\mathbf{x}})$ och då ger (10.3) oss att

$$\lim_{\mathbf{w} \rightarrow \bar{\mathbf{w}}} \frac{L(\mathbf{w}) - L(\bar{\mathbf{w}}) - (\mathbf{w} - \bar{\mathbf{w}})^t \beta(\bar{\mathbf{x}})}{\|\mathbf{w} - \bar{\mathbf{w}}\|} = 0$$

Alltså är L differentierbar i $\bar{\mathbf{w}}$ med gradient $\beta(\bar{\mathbf{x}})$. □

Vi vet nu, tack vare Thm 3.2.5 sid 86 i [5], att eftersom L är konkav så är den även subdifferentierbar, dvs den har subgradienter.

Sats 10.6.

Låt X vara en icke-tom, kompakt mängd i \mathbf{R}^n och låt $f : \mathbf{R}^n \rightarrow \mathbf{R}^m$ och $\beta : \mathbf{R}^n \rightarrow \mathbf{R}^{m+k}$ vara kontinuerliga så att för något $\bar{\mathbf{w}} \in \mathbf{R}^{m+k}$ är $X(\bar{\mathbf{w}})$ inte tom. Om $\bar{\mathbf{x}} \in X(\bar{\mathbf{w}})$ så är $\beta(\bar{\mathbf{x}})$ en subgradient till L i $\bar{\mathbf{w}}$

Bevis. Eftersom f , β kontinuerliga och X kompakt så $X(\bar{\mathbf{w}}) \neq \emptyset$ för något $\bar{\mathbf{w}} \in \mathbf{R}^{m+k}$. Låt $\bar{\mathbf{w}} \in \mathbf{R}^{m+k}$ och låt $\bar{\mathbf{x}} \in X(\bar{\mathbf{w}})$. Då

$$L(\mathbf{w}) = \inf \{f(\mathbf{x}) + \mathbf{w}^t \beta(\mathbf{x}); \mathbf{x} \in X\}$$

$$\leq f(\bar{\mathbf{x}}) + \mathbf{w}^t \beta(\bar{\mathbf{x}}) = f(\bar{\mathbf{x}}) + (\mathbf{w} - \bar{\mathbf{w}})^t \beta(\bar{\mathbf{x}}) + \bar{\mathbf{w}}^t \beta(\bar{\mathbf{x}}) = L(\bar{\mathbf{w}}) + (\mathbf{w} - \bar{\mathbf{w}})^t \beta(\bar{\mathbf{x}})$$

Alltså $L(\mathbf{w}) \leq L(\bar{\mathbf{w}}) + (\mathbf{w} - \bar{\mathbf{w}})^t \beta(\bar{\mathbf{x}})$ och $\beta(\bar{\mathbf{x}})$ är en subgradient till L i $\bar{\mathbf{w}}$. □

Definition 10.7. Riktningderivatan⁵, $f'(\bar{\mathbf{x}}; \mathbf{d})$, med riktning \mathbf{d} är, om det existerar, följande gränsvärde

$$f'(\bar{\mathbf{x}}; \mathbf{d}) = \lim_{\lambda \rightarrow 0^+} \frac{f(\bar{\mathbf{x}} + \lambda \mathbf{d}) - f(\bar{\mathbf{x}})}{\lambda}$$

För beviset för de nedanstående Sats 10.8 och Följdsats 10.9 hänvisar vi till sida 218 i [5].

Sats 10.8.

Låt X vara en icke-tom, kompakt mängd i \mathbf{R}^n och låt $f : \mathbf{R}^n \rightarrow \mathbf{R}^m$ och $\beta : \mathbf{R}^n \rightarrow \mathbf{R}^{m+k}$ vara kontinuerliga. Låt $\bar{\mathbf{w}}, \mathbf{d} \in \mathbf{R}^{m+k}$, då uppfyller riktningderivatan, i riktning \mathbf{d} , av L i punkten $\bar{\mathbf{w}}$ att

$$L'(\bar{\mathbf{w}}; \mathbf{d}) \geq \mathbf{d}^t \beta(\bar{\mathbf{x}}) \quad \text{för något } \bar{\mathbf{x}} \in X(\bar{\mathbf{w}})$$

Följdsats 10.9. Låt $\partial L(\bar{\mathbf{w}}) = \{\text{subgradienter till } L \text{ i punkten } \bar{\mathbf{w}}\}$. Antag att antagandena i Sats 10.8 håller. Då gäller

$$L(\bar{\mathbf{w}}; \mathbf{d}) = \inf \{\mathbf{d}^t \xi; \xi \in \partial L(\bar{\mathbf{w}})\}$$

⁵Fritt översatt från engelskans *directional derivative*

Sats 10.10.

Låt X vara en icke-tom, kompakt mängd i \mathbf{R}^n och låt $f : \mathbf{R}^n \rightarrow \mathbf{R}^m$ och $\beta : \mathbf{R}^n \rightarrow \mathbf{R}^{m+k}$ vara kontinuerliga. Då är ξ en subgradient till L i punkten $\bar{\mathbf{w}} \in \mathbf{R}^{m+k}$ om och endast om ξ tillhör det konvexa höljet av $\{\beta(\mathbf{y}); \mathbf{y} \in X(\bar{\mathbf{w}})\}$.

Bevis. Låt $\Gamma = \{\beta(\mathbf{y}); \mathbf{y} \in X(\bar{\mathbf{w}})\}$ och det konvexa höljet vara $H(\Gamma)$. Då gäller enligt Sats 10.6 att $\Gamma \subseteq \partial L(\bar{\mathbf{w}})$ och eftersom $\partial L(\bar{\mathbf{w}})$ är konvex så har vi att $H(\Gamma) \subseteq \partial L(\bar{\mathbf{w}})$. Eftersom X kompakt och β kontinuerlig så är även Γ kompakt (se [7]) och ett konvext hölje på en kompakt mängd är sluten. Alltså är $H(\Gamma)$ en sluten och kompakt mängd. Vi vill nu visa att $\partial L(\bar{\mathbf{w}}) \subseteq H(\Gamma)$. Vi gör ett motsägelsebevis.

Antag att det existerar $\xi' \in L(\bar{\mathbf{w}})$ men att $\xi' \notin H(\Gamma)$. Eftersom $H(\Gamma)$ är konvex och sluten så existerar det en skalär α och en vektor \mathbf{d} så att

$$\mathbf{d}^t \beta(\mathbf{y}) \geq \alpha \quad \text{för varje } \mathbf{y} \in X(\bar{\mathbf{w}}) \quad (10.4)$$

$$\mathbf{d}^t \xi' < \alpha \quad (10.5)$$

Enligt Sats 10.8 så existerar det $\mathbf{y} \in X(\bar{\mathbf{w}})$ så att $L'(\bar{\mathbf{w}} : \mathbf{d}) \geq \mathbf{d}^t \beta(\mathbf{y})$, det tillsammans med (10.4) så får vi att $L'(\bar{\mathbf{w}} : \mathbf{d}) \geq \alpha$. Men med Följdsats 10.9 och Sats 10.8 så har vi

$$L'(\bar{\mathbf{w}} : \mathbf{d}) = \inf \{ \mathbf{d}^t \xi; \xi \in L(\bar{\mathbf{w}}) \} \leq \mathbf{d}^t \xi' < \alpha$$

□

Exempel 8. Betrakta problemet

$$\text{Minimera } -(x_1 - 4)^2 - (x_2 - 4)^2$$

$$\begin{aligned} \text{Då} \quad & x_1 - 3 \leq 0 \\ & -x_1 + x_2 - 2 \leq 0 \\ & x_1 + x_2 - 4 \leq 0 \\ & x_1, x_2 \geq 0 \end{aligned}$$

Låt $g_1(x_1, x_2) = x_1 - 3$, $g_2(x_1, x_2) = -x_1 + x_2 - 2$ och

$X = \{(x_1, x_2) : x_1 + x_2 - 4 \leq 0; x_1, x_2 \geq 0\}$. Den duala funktionen är då

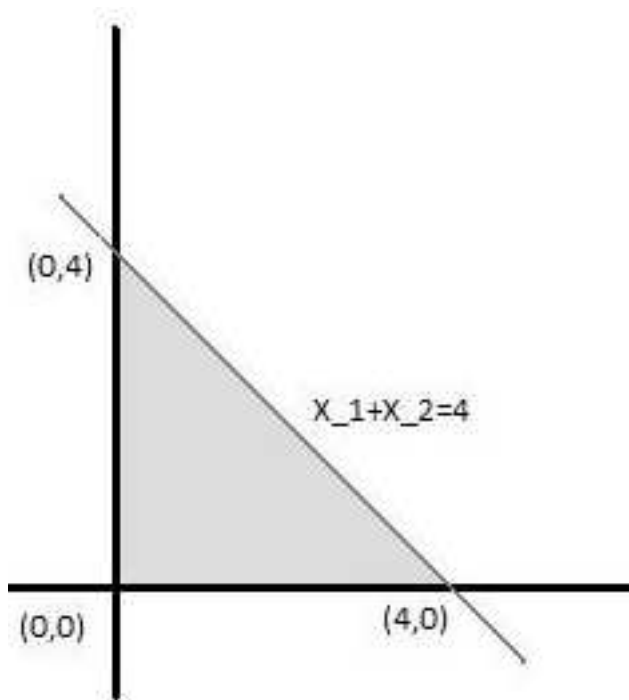
$$L(u_1, u_2) = \inf \{ -(x_1 - 4)^2 - (x_2 - 4)^2 + u_1(x_1 - 3) + u_2(-x_1 + x_2 - 2); \mathbf{x} \in X \}$$

Vi använder Sats 10.10 för att bestämma mängden av subgradients till L i punkten $\bar{\mathbf{u}} = (1, 5)^t$ (godtycklig punkt). För att hitta mängden $X(\bar{\mathbf{u}})$ behöver vi nu lösa följande problem:

$$\text{Minimera } -(x_1 - 4)^2 - (x_2 - 4)^2 - 4x_1 + 5x_2 - 13$$

$$\begin{aligned} \text{Då} \quad & x_1 + x_2 - 4 \leq 0 \\ & x_1, x_2 \geq 0 \end{aligned}$$

Där $f(x_1, x_2) = -(x_1 - 4)^2 - (x_2 - 4)^2 - 4x_1 + 5x_2 - 13$ är konkav eftersom $H(\mathbf{x})$ (den Hessianska matrisen) är negativt semidefinit. Detta ger att min $f(x_1, x_2)$ återfinns i någon av extrempunkterna $(0, 0)$, $(4, 0)$, $(0, 4)$. Se figur 1.



Figur 1: extrempunkterna

Då $f(0, 0) = f(4, 0) = -45$ och $f(0, 4) = -9$ så ser vi att de optimala lösningarna till delproblemet är $(0, 0)$ och $(4, 0)$, dvs $X(\bar{\mathbf{u}}) = \{(0, 0), (4, 0)\}$. Sats 10.10 ger att subgradienterna till L i $\bar{\mathbf{u}}$ ges av de konvexa kombinationerna av $\mathbf{g}(0, 0)$ och $\mathbf{g}(4, 0)$, vilket är detsamma som de konvexa kombinationerna av de två vektorerna $(-3, -2)^t$ och $(1, -6)^t$. \triangle

Med det duala problemet så vill vi maximera L under villkoret att $\mathbf{u} \geq \mathbf{0}$. Givet en punkt $\mathbf{w}^t = (\mathbf{u}^t, \mathbf{v}^t)$ så måste vi också undersöka i vilken riktning som L ökar.

Definition 10.11. En vektor \mathbf{d} kallas för *ascentriktning*⁶ av L , dvs i vilken riktning L ökar, i punkten \mathbf{w} om det existerar ett $\delta > 0$ så att

$$L(\mathbf{w} + \lambda \mathbf{d}) > L(\mathbf{w}) \quad \text{för varje } \lambda \in (0, \delta)$$

⁶Fritt översatt från engelskans *ascent direction*

Notera att om L är konkav så är \mathbf{d} en ascentriktning i \mathbf{w} om och endast om $L'(\mathbf{w}; \mathbf{d}) > 0$. L antar sitt maximum i \mathbf{w} om och endast om L inte har någon ascentriktning i \mathbf{w} , dvs om och endast om $L'(\mathbf{w}; \mathbf{d}) \leq 0$ för varje \mathbf{d} .

Från Sats 10.8 följer att vektor \mathbf{d} är en ascentriktning av L i \mathbf{w} om och endast om $\inf \{\mathbf{d}^t \xi : \xi \in \partial L(\mathbf{w})\} > 0$, alltså om och endast om följande olikhet håller för något $\varepsilon > 0$

$$\mathbf{d}^t \xi \geq \varepsilon > 0 \quad \text{för varje } \xi \in \partial L(\mathbf{w})$$

Eftersom vi vill maximera L så är vi inte bara intresserade av ascentriktning utan också av den riktning där L lokalt ökar som mest.

Definition 10.12. En vektor $\bar{\mathbf{d}}$ kallas för den *brantaste ascentriktningen* av L i \mathbf{w} om

$$L'(\mathbf{w}; \bar{\mathbf{d}}) = \max_{\|\mathbf{d}\| \leq 1} L'(\mathbf{w}; \mathbf{d})$$

Exempel 9.

$$\min \quad -2x_1 + 2x_2 + x_3 - 3x_4$$

$$\begin{aligned} \text{då} \quad & x_1 + x_2 + x_3 + x_4 \leq 8 \\ & x_1 - 2x_3 + x_4 \leq 2 \\ & x_1 + x_2 \leq 8 \\ & x_3 + 2x_4 \leq 6 \\ & x_1, x_2, x_3, x_4 \geq 0 \end{aligned}$$

Låt $X = \{(x_1, x_2, x_3, x_4 : x_1 + x_2 \leq 8, x_3 + 2x_4 \leq 6; x_1, x_2, x_3, x_4 \geq 0)\}$. Den duala funktionen blir då

$$\begin{aligned} \max \quad & L(u_1, u_2) \\ & u_1, u_2 \geq 0 \end{aligned}$$

Där $L(u_1, u_2) =$

$$\begin{aligned} &= \min\{-2x_1 + 2x_2 + x_3 - 3x_4 + u_1(x_1 + x_2 + x_3 + x_4 - 8) \\ &\quad + u_2(x_1 - 2x_3 + x_4 - 2) : x_1, x_2, x_3, x_4 \geq 0\} = \\ &= \underbrace{\min\{(-2 + u_1 + u_2)x_1 + (2 + u_1)x_2 ; x_1 + x_2 \leq 8, x_1, x_2 \geq 0\}}_{L_1(u)} + \\ &+ \underbrace{\min\{(1 + u_1 - 2u_2)x_3 + (-3 + u_1 + u_2)x_4\}; x_3 + 2x_4 \leq 6, x_3, x_4 \geq 0}_{L_2(u)} - 8u_1 - 2u_2 = \\ &=: L_1(u) + L_2(u) - 8u_1 - 2u_2 \quad \text{där } u_1, u_2 \geq 0 \end{aligned}$$

$$L_1(u) = \begin{cases} 0 & \text{om } u_1 + u_2 \geq 2 \\ 8(-2 + u_1 + u_2) & \text{om } u_1 + u_2 < 2 \end{cases}$$

och

$$L_2(u) = \begin{cases} 0 & \text{om } u_1 + u_2 > 3 \text{ och } u_1 - 2u_2 > -1 \text{ (I)} \\ 6(1 + u_1 + u_2) & \text{om } 1 + u_1 - 2u_2 < 0 < -3 + u_1 + u_2 \text{ eller (II)} \\ & -3 + u_1 + u_2 < 0, 1 + u_1 - 2u_2 < 0 \text{(III)} \\ 3(-3 + u_1 + u_2) & \text{om } -3 + u_1 + u_2 < 0 < 1 + u_1 - 2u_2 \text{ eller (IV)} \\ & -3 + u_1 + u_2 < 0, 1 + u_1 - 2u_2 < 0 \text{(V)} \end{cases}$$

$u_1, u_2 \geq 0$ för alla fall.

Eftersom $(4, 0) \in (I)$ inte är en punkt där funktionen ändrar utseende så är $L(u)$ differentierbar i $(4, 0)$. I området (I) är $L(u) = -8u_1 - 2u_2$ (ty $L_1 = L_2 = 0$).

$$\nabla L = (-8, -2) = -2(4, 1) \implies \nabla L(4, 0) = -2(4, 1)$$

$(4, 0) + \lambda \nabla L(4, 0) = (4 - 4\lambda, -\lambda)$ som inte ligger i det tillåtna området ty $-\lambda < 0$, alltså så är $\nabla L(4, 0)$ ej tillåten. Ta $d = (-7, 4)$, $(4, 0) + \lambda(-7, 4) = (4 - 7\lambda, 4\lambda) > 0$ för små $\lambda > 0$.

Eftersom L ska maximeras så letar vi efter en ascentriktning.

$$L((4, 0) + \lambda d) = -8(4 - 7\lambda) - 2 \cdot 4\lambda = -32 + 48\lambda > -32 = L(4, 0)$$

\implies att d är en ascentriktning enligt definitionen. \triangle

Sats 10.13 nedan visar att den brantaste ascentriktningen av Lagranges duala funktion ges av den subgradient som har den minsta normen.

Sats 10.13.

Låt X vara en icke-tom, kompakt mängd i \mathbf{R}^n och låt $f : \mathbf{R}^n \longrightarrow \mathbf{R}^m$ och $\beta : \mathbf{R}^n \longrightarrow \mathbf{R}^{m+k}$ vara kontinuerliga. Den brantaste ascentriktningen $\bar{\mathbf{d}}$ av L i \mathbf{w} är:

$$\bar{\mathbf{d}} = \begin{cases} \mathbf{0} & \text{om } \bar{\xi} = \mathbf{0} \\ \frac{\bar{\xi}}{\|\bar{\xi}\|} & \text{om } \bar{\xi} \neq \mathbf{0} \end{cases}$$

där $\bar{\xi}$ är den subgradient i $\partial L(\mathbf{w})$ som har den minsta normen.

Bevis. Enligt Definition 10.11 och Sats 10.8-Följdsats 10.9 så kan vi få den brantaste ascentriktningen från följande uttryck:

$$\max_{\|d\| \leq 1} L'(\mathbf{w}; \mathbf{d}) = \max_{\|d\| \leq 1} \inf \{ \mathbf{d}^t \xi : \xi \in \partial L(\mathbf{w}) \}$$

Vi vet också att

$$\begin{aligned} \max_{\|d\| \leq 1} L'(\mathbf{w}; \mathbf{d}) &= \max_{\|d\| \leq 1} \inf \{ \mathbf{d}^t \xi : \xi \in \partial L(\mathbf{w}) \} \\ &\leq \inf_{\xi \in \partial L(\mathbf{w})} \max_{\|d\| \leq 1} \mathbf{d}^t \xi = \inf_{\xi \in \partial L(\mathbf{w})} \|\xi\| = \|\bar{\xi}\| \end{aligned} \quad (10.6)$$

Om vi bildar en riktning $\bar{\mathbf{d}}$ så att $L'(\mathbf{w}; \bar{\mathbf{d}}) = \|\bar{\xi}\|$ så vet vi genom (10.6) att $\bar{\mathbf{d}}$ är den brantaste ascentriktningen. Om $\bar{\xi} = \mathbf{0}$ så gäller uppenbart för $\bar{\mathbf{d}} = \mathbf{0}$ att $L'(\mathbf{w}; \bar{\mathbf{d}}) = \|\bar{\xi}\|$. Antag nu att $\bar{\xi} \neq \mathbf{0}$ och låt $\bar{\mathbf{d}} = \frac{\bar{\xi}}{\|\bar{\xi}\|}$

$$\begin{aligned} L'(\mathbf{w}; \bar{\mathbf{d}}) &= \inf \{ \mathbf{d}^t \xi : \xi \in \partial L(\mathbf{w}) \} \\ &= \inf \left\{ \frac{\bar{\xi}^t \xi}{\|\bar{\xi}\|} : \xi \in \partial L(\mathbf{w}) \right\} \\ &= \frac{1}{\|\bar{\xi}\|} \inf \{ \|\bar{\xi}\|^2 + \bar{\xi}^t (\xi - \bar{\xi}) : \xi \in \partial L(\mathbf{w}) \} \\ &= \|\bar{\xi}\| + \frac{1}{\|\bar{\xi}\|} \inf \{ \bar{\xi}^t (\xi - \bar{\xi}) : \xi \in \partial L(\mathbf{w}) \} \end{aligned} \quad (10.7)$$

Eftersom $\|\bar{\xi}\|$ är den kortaste vektorn i $\partial L(\mathbf{w})$ så är $\bar{\xi}^t (\xi - \bar{\xi}) \geq 0$ för varje $\xi \in \partial L(\mathbf{w})$. Därmed är $\inf \{ \bar{\xi}^t (\xi - \bar{\xi}) : \xi \in \partial L(\mathbf{w}) \} = 0$ uppnådd i $\bar{\xi}$. Från (10.7) följer då att $L'(\mathbf{w}; \bar{\mathbf{d}}) = \|\bar{\xi}\|$ och vi har visat att vektorn $\bar{\mathbf{d}}$ är den brantaste ascentriktningen både när $\bar{\xi} = \mathbf{0}$ och när $\bar{\xi} \neq \mathbf{0}$. \square

11 Primala och duala problemen

11.1 Att formulera det duala problemet

Givet ett primalt problem (P)

$$\text{Minimera} \quad f(\mathbf{x})$$

$$\text{Då} \quad \begin{aligned} \mathbf{g}(\mathbf{x}) &\leq \mathbf{0} \\ \mathbf{h}(\mathbf{x}) &= \mathbf{0} \\ \mathbf{x} &\in X \end{aligned}$$

så har vi definierat Lagranges duala problem (D): $\max L(\mathbf{u}, \mathbf{v})$ under villkoret $\mathbf{u} \geq \mathbf{0}$. Där vi har uppskattat $L(\mathbf{u}, \mathbf{v})$ via Lagranges delproblem $L(\mathbf{u}, \mathbf{v}) = \min\{f(\mathbf{x}) + \mathbf{u}^t \mathbf{g}(\mathbf{x}) + \mathbf{v}^t \mathbf{h}(\mathbf{x}); \mathbf{x} \in X\}$ Vi har även beskrivit flera egenskaper hos den duala funktionen, speciellt så kräver det duala problemet maximeringen av en konkav funktion $L(\mathbf{u}, \mathbf{v})$ över den enkla begränsningsmängden $\{(\mathbf{u}, \mathbf{v}) : \mathbf{u} \geq \mathbf{0}\}$. Om L är differentierbar med egenskaperna som är angivna i Sats 10.5 så är $\nabla L(\bar{\mathbf{u}}, \bar{\mathbf{v}})^t = [\mathbf{g}(\bar{\mathbf{x}})^t, \mathbf{h}(\bar{\mathbf{x}})^t]$.

Det finns flera olika algoritmer för att maximera differentierbara konkava funktioner som man använder för att lösa det duala problemet, dessa algoritmer kräver en lämplig ascentriktning \mathbf{d} och en 1-dimensionell linje i denna riktning för att hitta en ny förbättrad lösning.

11.2 Skärningsplan- eller yttre linjäriseringsmetoden

Vi kommer nu att visa en metod för att lösa det duala problemet för vilken, i varje upprepning, en funktion som approximerar den duala funktionen $L(\mathbf{u}, \mathbf{v})$ är optimerad. Det finns flera sätt att lösa det duala problemet på men vi har tyvärr inte tid att i denna uppsats granska dessa närmare utan vi hänvisar till lämpliga böcker om optimering (se t.ex. referenslistan)

Låt $z = L(\mathbf{u}, \mathbf{v})$ då måste olikheten $z \leq f(\mathbf{x}) + \mathbf{u}^t \mathbf{g}(\mathbf{x}) + \mathbf{v}^t \mathbf{h}(\mathbf{x})$ hålla för alla $\mathbf{x} \in X$. Därför blir det duala problemet att maximera $L(\mathbf{u}, \mathbf{v})$ över $\mathbf{u} \geq \mathbf{0}$ ekvivalent med följande problem

$$\text{Maximera} \quad z$$

$$\text{Då} \quad z \leq f(\mathbf{x}) + \mathbf{u}^t \mathbf{g}(\mathbf{x}) + \mathbf{v}^t \mathbf{h}(\mathbf{x}) \quad \text{för } \mathbf{x} \in X \quad (11.1)$$

$$\mathbf{u} \geq \mathbf{0}$$

Detta är linjärt problem i variablerna z , \mathbf{u} och \mathbf{v} . Tyvärr så är bivillkoren oändliga och är inte kända exakt. Antag att vi har punkterna $\mathbf{x}_1, \dots, \mathbf{x}_{k-1}$ i X och studera istället följande approximerade problem:

Maximera z

$$\text{Då} \quad z \leq f(\mathbf{x}_j) + \mathbf{u}^t \mathbf{g}(\mathbf{x}_j) + \mathbf{v}^t \mathbf{h}(\mathbf{x}_j) \quad \text{för } j = 1, \dots, k-1 \quad (11.2)$$

$$\mathbf{u} \geq \mathbf{0}$$

Detta är ett linjärt problem med ändligt många bivillkor och kan lösas med hjälp av till exempel *Simplex-metoden*. Låt $L(z_k, \mathbf{u}_k, \mathbf{v}_k)$ vara en optimal lösning till det approximerade problemet, ibland kallat *master program*. Om denna lösning uppfyller (11.1) så är det lösning till Lagranges duala problem. För att kontrollera om (11.1) är uppfylld så tittar vi på följande delproblem:

$$\text{Minimera} \quad f(\mathbf{x}) + \mathbf{u}_k^t \mathbf{g}(\mathbf{x}) + \mathbf{v}_k^t \mathbf{h}(\mathbf{x})$$

$$\text{Då} \quad \mathbf{x} \in X$$

Låt \mathbf{x}_k vara en optimal lösning så att $L(\mathbf{u}_k, \mathbf{v}_k) = f(\mathbf{x}_k) + \mathbf{u}_k^t \mathbf{g}(\mathbf{x}_k) + \mathbf{v}_k^t \mathbf{h}(\mathbf{x}_k)$. Om $z_k \leq L(\mathbf{u}_k, \mathbf{v}_k)$ så är $(\mathbf{u}_k, \mathbf{v}_k)$ en optimal lösning till Lagranges duala problem, om $z > L(\mathbf{u}_k, \mathbf{v}_k)$ så är för $(\mathbf{u}, \mathbf{v}) = (\mathbf{u}_k, \mathbf{v}_k)$ olikheten i (11.1) inte uppfylld för $\mathbf{x} = \mathbf{x}_k$ och om så är fallet så lägger vi till villkoret

$$z \leq f(\mathbf{x}_k) + \mathbf{u}^t \mathbf{g}(\mathbf{x}_k) + \mathbf{v}^t \mathbf{h}(\mathbf{x}_k)$$

till bivillkoren i (11.2) och så löser vi det approximerade problemet igen. Den nuvarande optimala punkten $(z_k, \mathbf{u}_k, \mathbf{v}_k)$ motsäger uppenbart det villkor som vi har lagt till men denna punkt är borttagen därav namnet skärningsplansalgoritmen.

Skärningsplansmetoden (kort sammanfattning):

Antag att f , \mathbf{g} och \mathbf{h} är kontinuerliga och att X är kompakt så att mängden $X(\mathbf{u}, \mathbf{v})$ är icke-tom

Steg 1 Hitta en punkt $\mathbf{x}_0 \in X$ så att $\mathbf{g}(\mathbf{x}_0) \leq \mathbf{0}$ och $\mathbf{h}(\mathbf{x}_0) = \mathbf{0}$. Låt $k = 1$ och gå till huvudsteget.

Huvudsteg Lös följande approximerade problem (*master program*):

Maximera z

Då $z \leq f(\mathbf{x}_j) + \mathbf{u}^t \mathbf{g}(\mathbf{x}_j) + \mathbf{v}^t \mathbf{h}(\mathbf{x}_j)$ för $j = 1, \dots, k-1$

$\mathbf{u} \geq \mathbf{0}$

Låt $L(z_k, \mathbf{u}_k, \mathbf{v}_k)$ vara en optimal lösning och lös följande problem:

Minimera $f(\mathbf{x}) + \mathbf{u}_k^t \mathbf{g}(\mathbf{x}) + \mathbf{v}_k^t \mathbf{h}(\mathbf{x})$

Bivillkor $\mathbf{x} \in X$

Låt \mathbf{x}_k vara en optimal punkt och låt $L(\mathbf{u}_k, \mathbf{v}_k) = f(\mathbf{x}_k) + \mathbf{u}_k^t \mathbf{g}(\mathbf{x}_k) + \mathbf{v}_k^t \mathbf{h}(\mathbf{x}_k)$. Om $z_k \leq L(\mathbf{u}_k, \mathbf{v}_k)$ så stanna för då är $(\mathbf{u}_k, \mathbf{v}_k)$ en optimal dual lösning, men om $z > L(\mathbf{u}_k, \mathbf{v}_k)$ så lägg till villkoret $z \leq f(\mathbf{x}_k) + \mathbf{u}_k^t \mathbf{g}(\mathbf{x}_k) + \mathbf{v}_k^t \mathbf{h}(\mathbf{x}_k)$ till det approximerade problemet, ersätt k med $k+1$ och upprepa huvudsteget.

Vid varje upprepning då ytterligare ett bivillkor läggs till så ökar storleken på det approximerade problemet monotont. I praktiken så kan storleken på problemet bli väldigt stor, alla bivillkor som inte är bindande kan man ta bort. Teoretiskt så kanske inte detta garanterar konvergens. Man bör också notera att de optimala lösningarnas värden av det approximerade problemet bildar en icke-ökande följd $\{z_k\}$. Eftersom varje z_k är en övre begränsning till det optimala värdet för det duala problemet så kan vi stanna om $z_k - \max_{1 \leq j \leq k} L(\mathbf{u}_j, \mathbf{v}_j) < \varepsilon$, där ε är ett litet positivt tal.

Exempel 10.

Minimera $(x_1 - 2)^2 + \frac{1}{4}x_2^2$

Då $x_1 - \frac{7}{2}x_2 - 1 \leq 0$
 $2x_1 + 3x_2 = 4$

Vi låter $X = \{(x_1, x_2); 2x_1 + 3x_2 = 4\}$ så att Lagranges duala funktion blir

$$L(u) = \min\{(x_1 - 2)^2 + \frac{1}{4}x_2^2 + u(x_1 - \frac{7}{2}x_2 - 1); 2x_1 + 3x_2 = 4\} \quad (11.3)$$

Vi börjar med att hitta en möjlig lösning $\mathbf{x}_0 = (\frac{5}{4}, \frac{1}{2})^t$. Steg 1 är att lösa följande problem

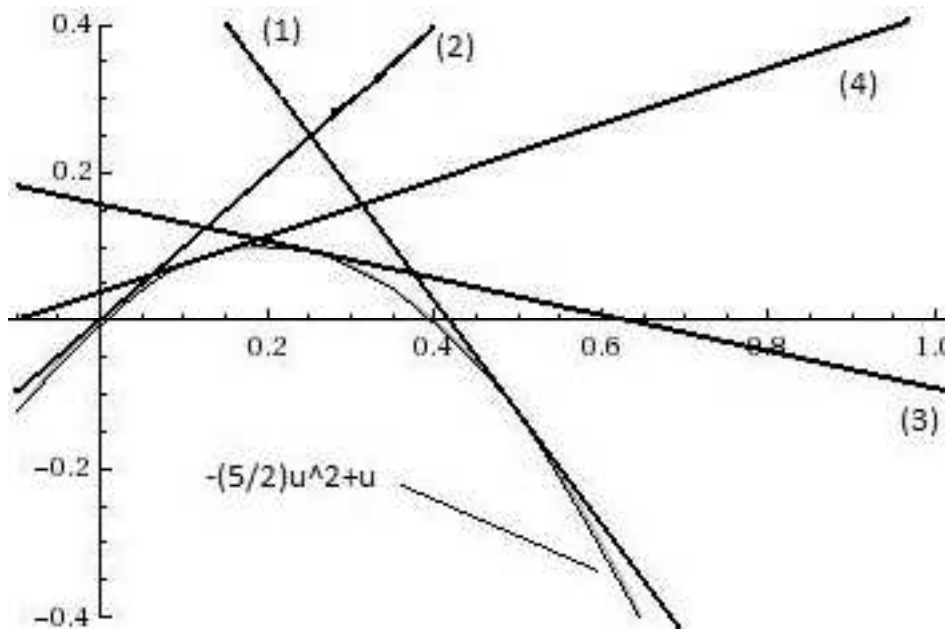
Maximera z

Då $z \leq \frac{5}{8} - \frac{3}{2}u$
 $u \geq 0$

Optimal lösning är $(z_1, u_1) = (\frac{5}{8}, 0)$. I steg 2 så löser vi (11.3) för $u = u_1 = 0$ och vi får den optimala lösningen $\mathbf{x}_1 = (2, 0)$ och $L(u_1) = 0 < \frac{5}{8} = z_1$ och vi ser att fler upprepningar är nödvändiga.

Upprepning	Tillagt villkor	Steg 1 lösning		Steg 2 lösning
		(z_k, u_k)	\mathbf{x}_k^t	$L(u_k)$
1	$z \leq \frac{5}{8} - \frac{3}{2}u$	$(\frac{5}{8}, 0)$	$(2, 0)$	0
2	$z \leq 0 + u$	$(\frac{1}{4}, \frac{1}{4})$	$(\frac{13}{8}, \frac{1}{4})$	$\frac{3}{32}$
3	$z \leq \frac{5}{32} - \frac{1}{4}u$	$(\frac{1}{8}, \frac{1}{8})$	$(\frac{29}{16}, \frac{1}{8})$	$\frac{11}{128}$
4	$z \leq \frac{5}{128} + \frac{3}{8}u$	$(\frac{7}{64}, \frac{3}{16})$	$(\frac{55}{32}, \frac{3}{16})$	$\frac{51}{512}$

Som vi ser i Figur 2 så är funktionen $L(u) = -\frac{5}{2}u + u$ och hyperplanen 1-4 är tangenter till L i punkterna (z_k, u_k) . Den duala målfunktionen är maximerad i $\bar{u} = \frac{1}{5}$ och $L(\frac{1}{5}) = \frac{1}{10} \triangle$



Figur 2:

11.3 Primala problemet

Vi har så här långt i denna uppsats lagt tyngden på det duala problemet men vårt mål är ju faktiskt att optimera det primala problemet och hitta en optimal lösning till målfunktionen.

Sats 11.1. Låt (\mathbf{u}, \mathbf{v}) vara en given vektor där $\mathbf{u} \geq \mathbf{0}$. Betrakta problemet att minimera $f(\mathbf{x}) + \mathbf{u}^t \mathbf{g}(\mathbf{x}) + \mathbf{v}^t \mathbf{h}(\mathbf{x})$ under villkoret $\mathbf{x} \in X$. Låt $\bar{\mathbf{x}}$ vara en optimal lösning till det problemet. Då är även $\bar{\mathbf{x}}$ en optimal lösning till följande problem:

Minimera $f(\mathbf{x})$

$$\begin{aligned} \text{Då} \quad & g_i(\mathbf{x}) \leq g_i(\bar{\mathbf{x}}) \quad \text{för } i \in I \\ & h_i(\mathbf{x}) = h_i(\bar{\mathbf{x}}) \quad \text{för } i = 1, \dots, l \\ & \mathbf{x} \in X \end{aligned}$$

där $I = \{i; u_i > 0\}$.

Bevis. Låt $\mathbf{x} \in X$ vara så att $h_i(\mathbf{x}) = h_i(\bar{\mathbf{x}})$ för $i = 1, \dots, l$ och $g_i(\mathbf{x}) \leq g_i(\bar{\mathbf{x}})$ för $i \in I$. Eftersom vårt problem är ett minimeringsproblem så har vi att:

$$f(\mathbf{x}) + \mathbf{u}^t \mathbf{g}(\mathbf{x}) + \mathbf{v}^t \mathbf{h}(\mathbf{x}) \geq f(\bar{\mathbf{x}}) + \mathbf{u}^t \mathbf{g}(\bar{\mathbf{x}}) + \mathbf{v}^t \mathbf{h}(\bar{\mathbf{x}}) \quad (11.4)$$

Men eftersom $\mathbf{h}(\mathbf{x}) = \mathbf{h}(\bar{\mathbf{x}})$ och $\mathbf{u}^t \mathbf{g}(\mathbf{x}) = \sum_{i \in I} u_i g_i(\mathbf{x}) \leq \sum_{i \in I} u_i g_i(\bar{\mathbf{x}}) = \mathbf{u}^t \mathbf{g}(\bar{\mathbf{x}})$ så får vi från (11.4) att

$$f(\mathbf{x}) + \mathbf{u}^t \mathbf{g}(\bar{\mathbf{x}}) \geq f(\mathbf{x}) + \mathbf{u}^t \mathbf{g}(\mathbf{x}) \geq f(\bar{\mathbf{x}}) + \mathbf{u}^t \mathbf{g}(\bar{\mathbf{x}})$$

vilket visar att $f(\mathbf{x}) \geq f(\bar{\mathbf{x}})$. Därför löser $\bar{\mathbf{x}}$ problemet givet i satsen. \square

12 Konjugatdualitet

Konjugat- eller Fenchedualitet som det också kallas baseras på en viktig tes om konvexitet, *konjugattransformationen*, vilken säger att till varje funktion f så associerar en konvex funktion, *konjugatet till f* .

$$\begin{aligned} \text{konjugatet till } f &= \text{konj } f \\ \text{konj}(\text{konj } f) &= f \end{aligned}$$

Det grundläggande sammanhanget för Fenchedualitet är problemet

$$\text{Minimera } f_1(x) - f_2(x) \tag{12.1}$$

$$\text{Då } x \in X_1 \cap X_2$$

där⁷ $f_1 : \mathfrak{R}^n \rightarrow \mathfrak{R}$ och $f_2 : \mathfrak{R}^n \rightarrow \mathfrak{R}$ är givna funktioner i \mathfrak{R}^n och X_1 och X_2 är givna delmängder av \mathfrak{R}^n .

Fenchedualitet är inte fundamentalt olik Lagrangedualitet som vi tittat på i tidigare kapitel. Vi kan i själva verket härleda Fenchedualitet från Lagrangedualiteten genom att konvertera (12.1) till följande ekvivalenta problem i variablerna $y, z \in \mathfrak{R}^n$

$$\begin{aligned} \text{Minimera } f_1(y) - f_2(z) \\ \text{Då } z = y \\ y \in X_1 \\ z \in X_2 \end{aligned} \tag{12.2}$$

Genom att dualisera villkoret $z = y$ så är den duala funktionen

$$\begin{aligned} q(\lambda) &= \inf \{f_1(y) - f_2(z) + (z - y)^t \lambda\} = \inf \{z^t \lambda - f_2(z)\} - \inf \{f_1(y) - y^t \lambda\} \\ &= f_2^*(\lambda) - f_1^*(\lambda) \end{aligned}$$

Där $(z - y)^t$ står för transponaten av $(x - y)$ och f_1^* och f_2^* är definierade som

$$f_1^*(\lambda) = \sup \{x^t \lambda - f_1(x)\}, \quad f_2^*(\lambda) = \inf \{x^t \lambda - f_2(x)\}$$

Duala problemet (D)

$$\text{Maximera } f_2^*(\lambda) - f_1^*(\lambda) \tag{12.3}$$

$$\text{Då } \lambda \in \Gamma_1 \cup \Gamma_2$$

⁷ $\mathfrak{R} = \mathbf{R} \setminus \{-\infty, \infty\}$

Där $\Gamma_1 = \{\lambda \mid f_1^*(\lambda) < \infty\}$ och $\Gamma_2 = \{\lambda \mid f_2^*(\lambda) > -\infty\}$. Om vi tittar på f_1^* och f_2^* i (12.3) som utvidgade reellvärda funktioner $f_1^* : \mathfrak{R}^n \rightarrow (-\infty, \infty]$ och $f_2^* : \mathfrak{R}^n \rightarrow [-\infty, \infty)$ så behöver inte Γ_1 och Γ_2 framstå så tydligt. Vi kan då uttrycka det duala problemet som

$$\text{Maximera } f_2^*(\lambda) - f_1^*(\lambda)$$

$$\text{Då } \lambda \in \mathfrak{R}^n$$

På samma sätt kan vi se på f_1 och f_2 som utvidgade reellvärda funktioner där de antar värdena ∞ respektive $-\infty$ då $x \notin X_1$ respektive $x \notin X_2$. Vi kan nu skriva det primala problemet som

$$\text{Minimera } f_1(y) - f_2(z) \tag{12.4}$$

$$\text{Då } x \in \mathfrak{R}^n$$

Där $f_1 : \mathfrak{R}^n \rightarrow (-\infty, \infty]$ och $f_2 : \mathfrak{R}^n \rightarrow [-\infty, \infty)$ är givna funktioner och f_1^* och f_2^* definieras som

$$f_1^*(\lambda) = \sup_{x \in \mathfrak{R}^n} \{x^t \lambda - f_1(x)\} \text{ och } f_2^*(\lambda) = \inf_{x \in \mathfrak{R}^n} \{x^t \lambda - f_2(x)\} \tag{12.5}$$

Funktionen f_1^* i (12.5) kallas för den *konvexa konjugatfunktionen* till f_1 och f_2^* kallas den *konkava konjugatfunktionen* till f_2 .

12.1 Konjugerade funktioner

Definition 12.1. Om $f : X \rightarrow [-\infty, \infty]$ är en given funktion på en godtycklig mängd $X \subset \mathfrak{R}^n$ så är dess *epigraf* mängden

$$\text{epi } f = \{(x, w) \mid x \in X, w \in \mathfrak{R}, f(x) \leq w\} \tag{12.6}$$

Definition 12.2. Om $f : X \rightarrow [-\infty, \infty]$ är en given funktion definierad i en mängd $X \subset \mathfrak{R}^n$ så är den *konvex* precis när dess epigraf är en konvex mängd i \mathfrak{R}^{n+1} .

Vi tittar nu på det konvexa konjugatet f^* till f som vi introducerade tidigare:

$$f^*(\lambda) = \sup_{x \in \mathfrak{R}^n} \{x^t \lambda - f(x)\} \quad \lambda \in \mathfrak{R}^n$$

Notera att vi utan att veta något om f har att f^* är en sluten konvex funktion. Vi har här inga krav på att f^* är en proper funktion även om f är det. Vi kommer dock att i nästa sats att visa att om f är konvex så är

f^* proper om och endast om f är proper. Vi kallar en funktion f *proper* om $f(x) < \infty$ för åtminstone ett $x \in X$ och $f(x) > -\infty$ för alla $x \in X$. Man kan med andra ord säga att en funktion är proper om och endast om dess epigraf är icke-tom och inte innehåller en vertikal linje. Om f inte är proper så säger vi att den är improper.

För beviset till nästa sats så hänvisar vi till [6].

Sats 12.3.

Låt $f : \mathbb{R}^n \rightarrow [-\infty, \infty]$ vara en funktion, låt f^* vara f 's konvexa konjugat och betrakta konjugatet av f^*

$$\tilde{f}(x) = \sup_{\lambda \in \mathbb{R}^n} \{\lambda^t x - f^*(\lambda)\} \quad x \in X$$

(i) Vi har

$$f(x) \geq \tilde{f}(x)$$

(ii) Om f är konvex och någon av funktionerna f , f^* och \tilde{f} är proper så ger det att även de andra två funktionerna är propra funktioner.

(iii) Om f är sluten, proper och konvex så är

$$f(x) = \tilde{f}(x) \quad \forall x \in \mathbb{R}^n$$

(iv) Låt \hat{f} vara det konvexa slutna höljet av f . Om \hat{f} uppfyller att $\hat{f} > -\infty$ $\forall x \in \mathbb{R}^n$ så är

$$\hat{f}(x) = \tilde{f}(x) \quad \forall x \in \mathbb{R}^n$$

Notera att antagandet om att $\hat{f}(x) > -\infty \forall x$ är viktigt för att (iv) ska hålla, här är ett exempel.

Exempel 11. Studera funktionen $f : \mathbb{R} \rightarrow [-\infty, \infty]$

$$f(x) = \begin{cases} \ln x & x > 0 \\ \infty & x \leq 0 \end{cases}$$

Vi ser att det konvexa slutna höljet av f ges av

$$\hat{f}(x) = \begin{cases} -\infty & x \geq 0 \\ \infty & x < 0 \end{cases}$$

Konjugatet till f ges av

$$f^*(\lambda) = \infty \quad \forall \lambda \in \mathfrak{R}^n$$

och vi har att konjugatet av f^* är

$$\tilde{f}(x) = \sup_{\lambda \in \mathfrak{R}^n} \{\lambda^t x - f^*(\lambda)\} = -\infty$$

Då konjugatet av $f^*(\lambda)$ antar värdet $-\infty$ överallt så är $\tilde{f}(x) \neq \hat{f}(x) \triangle$

12.2 Fenchels dualitetssatser

Studera problemet

$$\text{Maximera } f_1(x) - f_2(Qx) \tag{12.7}$$

$$\text{Då } x \in \mathfrak{R}^n$$

Där Q är en $m \times n$ -matris och $f_1 : \mathfrak{R}^n \rightarrow (-\infty, \infty]$ och $f_2 : \mathfrak{R}^m \rightarrow [-\infty, \infty)$ är utvidgade reellvärda funktioner. Som vi sagt tidigare så kan vi applicera Lagrangedualitet på (12.7) om problemet är skrivet i termer av f_1 och $-f_2$:s område på följande sätt

$$\text{Maximera } f_1(y) - f_2(z) \tag{12.8}$$

$$\text{Då } z - Qy = 0 \quad y \in \text{dom}(f_1) \quad z \in \text{dom}(-f_2)$$

Där $\text{dom}(f) = \{x \in X; f(x) < \infty\}$. Genom att dualisera likheten $z - Qy = 0$ så får vi den duala funktionen

$$\begin{aligned} q(\lambda) &= \inf_{y \in \mathfrak{R}^n, z \in \mathfrak{R}^m} \{f_1(y) - f_2(z) + (z - Qy)^t \lambda\} = \\ &= \inf_{z \in \mathfrak{R}^m} \{z^t \lambda - f_2(z)\} + \inf_{y \in \mathfrak{R}^n} \{f_1(y) - y^t Q^t \lambda\} = \\ &= f_2^*(\lambda) - f_1^*(Q^t \lambda) \end{aligned}$$

där $f_1^* : \mathfrak{R}^n \rightarrow (-\infty, \infty]$ och $f_2^* : \mathfrak{R}^m \rightarrow [-\infty, \infty)$ är

$$f_1^*(\lambda) = \sup_{x \in \mathfrak{R}^n} \{x^t \lambda - f_1(x)\} \quad f_2^*(\lambda) = \inf_{z \in \mathfrak{R}^m} \{z^t \lambda - f_2(z)\}$$

Det duala problemet är då

$$\text{Maximera } f_2^*(\lambda) - f_1^*(Q^t \lambda) \tag{12.9}$$

$$\text{Då } \lambda \in \mathfrak{R}^m$$

Sats 12.4. Fenchels primala sats

Låt funktionerna f_1 och $-f_2$ vara propra och konvexa. Då har vi att

$$\inf_{x \in \mathbb{R}^n} \{f_1(x_1) - f_2(Qx)\} = \sup_{x \in \mathbb{R}^n} \{f_2^*(\lambda) - f_1^*(Q^t\lambda)\}$$

och supremumet på högra sidan nås om följande villkor håller:

- (i) $\text{dom}(f_1)$ är skärningspunkten mellan en polyeder P_1 och en konvex mängd C_1 och f_1 kan utvidgas till en reellvärd konvex funktion på C_1 . (Alltså det existerar en konvex funktion $\bar{f}_1 : C_1 \rightarrow \mathbb{R}$ så att $f_1(x) = \bar{f}_1(x)$ för alla $x \in \text{dom}(f_1)$)
- (ii) $\text{dom}(-f_2)$ är skärningspunkten mellan en polyeder P_2 och en konvex mängd C_2 och f_2 kan utvidgas till en reellvärd konkav funktion på C_2 .
- (iii) $\{Q(\text{dom}(f_1) \cap \text{ri}(C_1))\} \cap \{\text{dom}(-f_2) \cap \text{ri}(C_2)\} \neq \emptyset$

Bevis. Villkor (iii) \implies

$$f^* = \inf \{f_1(x_1) - f_2(Qx); x \in \mathbb{R}^n\} < \infty$$

Om f^* är ändlig så håller denna sats genom tidigare resultat (Sats 8.13-Sats 9.2). Om $f^* = -\infty$ så är även $\sup \{g_2(\lambda) - g_1(Q^t\lambda)\} = -\infty$ (enligt Sats 8.1) för alla λ . \square

Om vi antar att funktionerna f_1 och $-f_2$ är slutna, propra och konvexa så visar Sats 12.3 att den konjugatet av den konvexa konjugatfunktionen till f_1^* är f_1 och att den konkava konjugatfunktionen till f_2^* är f_2 . Under dessa omständigheter är dualiteten symmetrisk, dvs det duala problemet är av samma typ som det primala och genom att dualisera det duala problemet så får vi det primala problemet. Följdaktligen så får vi, genom att applicera Sats 12.4 med x och λ omvända, följande:

Sats 12.5. Fenchels duala sats

Låt funktionerna f_1 och $-f_2$ vara slutna, propra och konvexa. Då har vi att

$$\inf_{x \in \mathbb{R}^n} \{f_1(x_1) - f_2(Qx)\} = \sup_{x \in \mathbb{R}^n} \{f_2^*(\lambda) - f_1^*(Q^t\lambda)\}$$

och infimumet på den vänstra sidan nås om följande villkor håller:

- (i) $\text{dom}(f_1^*)$ är skärningspunkten mellan en polyeder P_1 och en konvex mängd C_1 och g_1 kan utvidgas till en reellvärd konvex funktion på C_1 .
- (ii) $\text{dom}(-f_2^*)$ är skärningspunkten mellan en polyeder P_2 och en konvex mängd C_2 och g_2 kan utvidgas till en reellvärd konkav funktion på C_2 .
- (iii) $\{\text{dom}(f_1^*) \cap \text{ri}(C_1)\} \cap \{Q^t(\text{dom}(-f_2^*) \cap \text{ri}(C_2))\} \neq \emptyset$

13 Linjär-kvadratisk kontroll och dualitet

Vi ska i detta avsnitt försöka att beskriva sambandet mellan de två till synes helt skilda ämnena linjär-kvadratisk kontroll och dualitet och hur dessa är sammanbundna via konvex konjugatdualitet.

En speciell form av *linjär-kvadratisk* (LQ)⁸ kontrollteori är den *linjär-kvadratiske regleraren* (LQR)⁹ som är formulerad enligt följande:

$$\min x_T^t S x_T + \int_{\tau}^T (y^t(t) Q(t) y(t) + u^t(t) R u) dt$$

över alla integrerbara kontrollfunktioner $u : [\tau, T] \rightarrow \mathbf{R}^m$, under de linjära dynamiska villkoren

$$\dot{x}(t) = Ax(t) + Bu(t), \quad y(t) = Cx(t)$$

med det initiala värdet $x(\tau) = x_0$.

$x \in \mathbf{R}^n$ är tillståndsvektorn¹⁰, A , B , C , Q , R och S är vektorer av lämplig storlek. S , Q och R är symmetriska och positivt definita.

Standardmetoden för att lösa sådana här problem är att använda dynamisk programmering eller genom att använda Pontryagins maximeringsprincip, där lösningen nås via en Riccati-ekvation under lämpliga antaganden. Vi visar inte här hur man gör detta utan vi hänvisar till litteratur om optimal kontrollteori, t ex [9].

Konjugatfunktioner har vi tidigare definierat på följande sätt:

Låt $f : \mathfrak{R}^n \rightarrow [-\infty, \infty]$, där f är en konvex och proper funktion, då är f :s konvexa konjugatfunktion $f^* : \mathfrak{R}^n \rightarrow [-\infty, \infty]$ definieras av

$$f^*(\lambda) = \sup_{x \in \mathfrak{R}^n} \{x^t \lambda - f(x)\}$$

Vi vet att f^* är en konvex och proper funktion sedan tidigare resultat, vi vet även att konjugatet av f^* är f . Eftersom f och f^* är i samma funktionsrum så innebär det att, med dessa villkor, att den linjär-kvadratiske regleraren och dualen av den linjär-kvadratiske regleraren är symmetriska.

Ett enkelt exempel på ett par konjugatfunktioner är:

$$f(x) = \frac{1}{2} x^t Q x, \quad f^*(\lambda) = \frac{1}{2} \lambda^t Q^{-1} \lambda$$

⁸Linear quadratic

⁹Linear quadratic regulator

¹⁰Från engelskans *state vector*

där Q är symmetrisk och positivt definit matris vilket innebär att $-f(x)$ är konvex och har sitt minimum i $x = Q^{-1}\lambda$ och vi får då att

$$f^*(\lambda) = \lambda'Q^{-1}\lambda - \frac{1}{2}x^tQx = \lambda^tQ^{-1}\lambda - \frac{1}{2}\lambda^tQ^{-1}\lambda = \frac{1}{2}\lambda^tQ^{-1}\lambda$$

Ett annat exempel för samma Q och en sluten konvex icke-tom delmängd $X \subseteq \mathfrak{R}^n$ är:

$$f(x) = \begin{cases} \frac{1}{2}x^tQx & x \in X \\ \infty & x \notin X \end{cases}$$

$$f^*(\lambda) = \sup_{x \in X} \left\{ \lambda^t x - \frac{1}{2}x^tQx \right\}$$

Om $f(0) = 0$ så $0 \in X$, om $0 \in \text{int}X$ så är f^* kvadratisk på något område runt 0 och ges av $\frac{1}{2}\lambda^tQ^{-1}\lambda$.

Generellt så är f^* differentierbar och ∇f^* är *Lipschitz kontinuerlig*. En fundamental egenskap för f och f^* är att för något $x, \lambda \in \mathfrak{R}^n$ så:

$$f(x) + f^*(\lambda) \geq x'\lambda \tag{13.1}$$

och $f(x) = f^*(\lambda) = x^t\lambda$ om och endast om $\lambda \in \partial f(x)$ och om och endast om $\lambda \in \partial f^*(x)$, där ∂f är subdifferentialen av den konvexa funktionen f , se Definition 10.3. Detta är basen för Euler-Lagrange och Hamiltonian optimalitets villkor för linjär-kvadratiska kontrollproblem. För detaljer om ovanstående resultat så hänvisar vi till [8].

Generella dualitetsramar:

Givet konvexa undre semikontinuerliga funktioner $L, l : \mathfrak{R}^{2n} \rightarrow [-\infty, \infty]$ betrakta

$$\tilde{L}(p, w) = L^*(w, p) \quad \tilde{l}(p_\tau, p_T) = l^*(p_\tau, -p_T)$$

och det primala problemet (P):

$$\min_x \int_\tau^T L(x(t), \dot{x}(t))dt + l(x(\tau), x(T)) \tag{13.2}$$

och det duala problemet (D):

$$\min_p \int_\tau^T \tilde{L}(p(t), \dot{p}(t))dt + \tilde{l}(p(\tau), p(T)) \tag{13.3}$$

Notera att dualen av (D) är (P).

Från (13.1) följer att $\inf(P) \geq -\inf(D)$ och att p ger den undre begränsningen till $\inf(P)$ genom (13.3) medan x ger en undre begränsning till $\inf(D)$ genom (13.2). I [10] visar författarna att under några milda antaganden så är

$$-\infty < \inf(P) = -\inf(D) < \infty \quad (13.4)$$

och dessutom att de optimala lösningarna för det primala och duala problemet existerar.

Det linjär-kvadratiske regulatorproblemet kan skrivas om under det ramverk vi beskrivit ovan. För varje fixt $\tau \leq T$, $\xi \in \mathfrak{R}^n$ betraktar vi

$$L(x, v) = x^t C^t Q C x + \min\{u^t R u : Ax + Bu = v\}$$

$$l(x_\tau, x_T) = \delta_\xi(x_\tau) + x_T^t S x_T$$

där δ_ξ är indikatorfunktionen av ξ : $\delta_\xi(\xi) = 0$ och $\delta_\xi(x) = \infty$ om $x \neq \xi$. Om en lösning $x : [\tau, T] \rightarrow \mathfrak{R}^n$, till det resulterande primala problemet (P), betecknad som $P(\tau, \xi)$, för varje fixt τ och ξ , hittas så kan en optimal kontroll $u : [\tau, T] \rightarrow \mathfrak{R}^m$ för den linjär-kvadratiske regulatorn återfås, i nästan alla $t \in [\tau, T]$, som minimeringen till $u^t R u$ över alla u sådana att $Ax(t) + Bu = \dot{x}(t)$. Givet ett sådant L och l så är dualen av $P(\tau, \xi)$, betecknad som $D(\tau, \xi)$, definierad som

$$\tilde{L}(p, w) = p^t B R B^t p + \min_z \{z^t C Q C^t z : -A^t p + C^t z = w\}$$

$$\tilde{l}(p_\tau, p_T) = \xi^t p_\tau + p_T^t S p_T$$

Med L och l ovan så kan man sedan visa att (13.4) håller.

14 Stokastisk kontroll

I det här kapitlet fokuserar vi kort på den klassiska kontrolltekniken via Bellmans ekvation applicerat på en stokastisk miljö. Vi utgår från att läsaren är familjär med resultatet för deterministisk kalkyl och fokuserar endast på det stokastiska fallet. Vi utgår från en stokastisk differentialekvation¹¹

$$dX_t = f(X_t, U_t)dt + g(X_t)dB_t, \quad X_0 = x, \quad t \in [0, T] \quad (14.1)$$

Och söker en optimal lösning U_0 på en kostnadsfunktion C :

$$C(U) = E\left[\int_0^T F(X_\tau, U_\tau, \tau)d\tau\right]$$

Vi ställer nu upp Bellmans ekvation (som vi kallar för D för att undvika konflikt med notationen B för den browniska rörelsen) i punkten (x, t) :

$$\frac{\partial D}{\partial t} + \frac{1}{2}(g(x))^2 \frac{\partial^2 D}{\partial x^2} + \min_u \left[\frac{\partial}{\partial x} f(x, u) + F(x, u, t) \right] = 0 \quad (14.2)$$

som för alla x uppfyller $D(x, T) = 0$.

Nu tar vi några (X, U) som uppfyller (14.1) och får därför:

$$\begin{aligned} & \int_0^T F(X_t, U_t, t)dt = \\ &= \int_0^T \frac{\partial D}{\partial x}(X_t, t) f(X_t, U_t) + F(X_t, U_t, t)dt - \int_0^T \frac{\partial D}{\partial x}(X_t, t) f(X_t, U_t)dt \geq \\ &\geq \int_0^T \min \left[\frac{\partial D}{\partial x}(X_t, t) f(X_t, U_t) + F(X_t, U_t, t) \right] dt - \int_0^T \frac{\partial D}{\partial x}(X_t, t) f(X_t, U_t)dt \end{aligned}$$

Nu kan vi låna ett uttryck för $\min[\dots]$ ur (14.2) och därför får vi ovanstående till

$$= - \int_0^T \frac{\partial D}{\partial t} + \frac{1}{2}(g(x))^2 \frac{\partial^2 D}{\partial x^2} + \frac{\partial D}{\partial x} f(X_t, U_t)dt$$

För enklare beräkningar kallar vi funktionen under integralen för $H(t)$. Nu ger Sats 5.3:

$$d(D(X_t, t)) = H(t)dt + \frac{\partial D}{\partial x}(X_t, t)g(X_t)dB_t$$

¹¹I detta kapitel låter vi $X(t) = X_t$ för smidigare notation

Så alltså

$$\begin{aligned} \underbrace{D(x, T) - D(x, 0)}_{=0} &= \int_0^T H(t)dt + \int_0^T \frac{\partial D}{\partial x}(X_t, t)g(X_t)dB_t \Leftrightarrow \\ \Leftrightarrow D(x, 0) &= -\left(\int_0^T H(t)dt + \int_0^T \frac{\partial D}{\partial x}(X_t, t)g(X_t)dB_t\right) \Rightarrow \\ &\Rightarrow D(x, 0) = -E\left[\int_0^T H(t)dt\right] \end{aligned}$$

Och vi har visat att

$$C(U) = E[F(X_t, U_t, t)dt] \geq D(x, 0) \quad (14.3)$$

Nu behöver vi alltså finna ett U_0 sådant att $C(U_0) = D(x, 0)$.

Vi inleder med att definiera funktionen $\Phi(x, t)$ sådan att $\forall (x, t)$:

$$\frac{\partial D}{\partial x}f(x, \Phi(x, t)) + F(x, \Phi(x, t), t) = \min_u \left(\frac{\partial D}{\partial x}f(x, u) + F(x, u, t) \right) \quad (14.4)$$

Sedan löser vi (14.1) för $X_{0,t}$ och $\Phi(X_{0,t}, t)$ och definerar

$$U_{0,t} = \Phi(X_{0,t}, t)$$

Nu kan vi undersöka:

$$\begin{aligned} &\int_0^T F(X_0, U_0, t)dt = \\ &= \int_0^T \underbrace{\frac{\partial D}{\partial x}f(X_0, \Phi(X_0, t)) + F(X_0, \Phi(X_0, t), t)}_{=\min_u, (14.4)} dt - \\ &\quad - \int_0^T \frac{\partial D}{\partial x}f(X_0, \Phi(X_0, t))dt = \\ &= \int_0^T \underbrace{\frac{\partial D}{\partial t} + \frac{1}{2}(g(X_0))^2 \frac{\partial^2 D}{\partial x^2} + \frac{\partial D}{\partial x}f(X_0, \Phi(X_0, t))}_{=K_t} dt \end{aligned}$$

Sats 5.3 ger:

$$d(D(X_0, t)) = K_t + \frac{\partial D}{\partial x}g(X_0)dB$$

Och vi undersöker

$$0 - D(x, 0) = D(X_0(T), T) - D(X_0(0), 0) = \int_0^T K_t dt + \int_0^T \frac{\partial D}{\partial x}g(X_0(t))dB$$

Så om vi tar väntevärden av ovanstående får vi.

$$D(x, 0) = -E\left[\int_0^T K_t dt\right]$$

Som visar att:

$$C(U_0) = E\left[\int_0^T F(X_0, U_0, t) dt\right] = D(x, 0) \quad (14.5)$$

(14.3) och (14.5) visar att U_0 är en optimal kontroll.

15 Tillämpningar av icke-linjära problem

15.1 Stokastisk tillgångsallokering

Betrakta följande linjära programmeringsproblem:

$$\text{Maximera} \quad \mathbf{c}^t \mathbf{x}$$

$$\text{Då} \quad \begin{aligned} \mathbf{A} \mathbf{x} &\leq \mathbf{b} \\ x &\geq 0 \end{aligned}$$

\mathbf{A} är en $m \times n$ -matris där \mathbf{a}_i är den i :te kolumnvektorn, $\mathbf{c}_i \mathbf{x}$ är en n -vektor och \mathbf{b} är en m -vektor. Antag att vi har m stycken tillgångar representerade av vektor \mathbf{b} . Kolumn \mathbf{a}_j i \mathbf{A} är då aktivitet j och variabeln x_j representerar nivån på den valda aktiviteten.

Aktivitet j på nivån x_j använder $\mathbf{a}_j x_j$ av den totala tillgängliga tillgången $\mathbf{A} \mathbf{x} = \sum_{j=1}^n \mathbf{a}_j x_j$. Om vinsten av aktivitet j är c_j så är den totala nyttan $\sum_{j=1}^n c_j x_j = \mathbf{c}^t \mathbf{x}$. Problemet kan nu tolkas som att hitta det bästa sättet att allokera tillgångsvektorn \mathbf{b} med varierande möjliga aktiviteter så att nyttan maximeras.

Modellen här ovan är inte riktigt tillfredsställande då det i verkligheten inte är så enkelt ty nyttokoefficienterna c_1, c_2, \dots, c_n är i det verkliga fallet slumpmässiga och inte fixa. Därför kommer vi från och med nu anta att \mathbf{c} är en slumpmässig vektor med medelvärde $\bar{\mathbf{c}} = (c_1, \dots, c_n)$ och kovariansmatris \mathbf{V} . Den objektiva funktionen z kommer därför att vara en slumpmässig variabel med medelvärde $\bar{\mathbf{c}}^t \mathbf{x}$ och varians $\mathbf{x}^t \mathbf{V} \mathbf{x}$.

I ett realistiskt problem vill man ha så stort väntevärde som möjligt samtidigt som man vill ha en liten varians. Det finns olika sätt att simultant betrakta $E(z)$ och $\text{Var}(z)$. Antag att vi är intresserade av att hålla väntevärdet på en nivå så att det åtminstone är lika med ett värde \bar{z} , vilket man ofta kallar den eftersträfvade- eller tillfredsställande nivån. Vi kan då ställa upp problemet på följande sätt

$$\text{Minimera} \quad \mathbf{x}^t \mathbf{V} \mathbf{x}$$

$$\text{Då} \quad \begin{aligned} \mathbf{A} \mathbf{x} &\leq \mathbf{b} \\ \bar{\mathbf{c}}^t \mathbf{x} &\geq \bar{z} \\ \mathbf{x} &\geq \mathbf{0} \end{aligned}$$

Ett alternativt sätt är att låta $\alpha = P(\bar{\mathbf{c}}^t \mathbf{x} \geq \bar{z})$, uppenbart så vill vi maximera α . Antag att vektorn av slumpmässiga variabler \mathbf{c} kan uttryckas

som funktionen $\mathbf{d} + y\mathbf{f}$, där \mathbf{d} och \mathbf{f} är fixa vektorer och y är en slumpvariabel

$$\implies \alpha = P(\mathbf{d}^t \mathbf{x} + y \mathbf{f}^t \mathbf{x} \geq \bar{z}) = P\left(y \geq \frac{\bar{z} - \mathbf{d}^t \mathbf{x}}{\mathbf{f}^t \mathbf{x}}\right) \quad \text{om } \mathbf{f}^t \mathbf{x} > 0$$

Eftersom vi vill maximera α så får vi

$$\text{Minimera } \frac{\bar{z} - \mathbf{d}^t \mathbf{x}}{\mathbf{f}^t \mathbf{x}}$$

$$\text{Då } \begin{aligned} \mathbf{A}\mathbf{x} &\leq \mathbf{b} \\ \mathbf{x} &\geq 0 \end{aligned}$$

Riskaversions modellen

Inte någon av metoderna ovan har tagit hänsyn till att olika individer har olika riskbenägenheter, dvs de har olika riskbeteenden. En individ kan vara helt riskavert eller helt risktagande och allt däremellan.

För de flesta individer så sjunker värdet på exempelvis 100 kr när deras totala nettoförmögenhet ökar. Sambandet mellan värdet och ens egen förmögenhet z kallas *nyttan av* z , $u(z)$. Oftast, för enkelhetens skull, så brukar man normalisera den totala nyttan u så att $u = 0$ när $z = 0$ och $u = 1$ när $z \rightarrow \infty$, u är individens nyttofunktion och är vanligtvis inte konkav. Matematiskt kan nyttofunktionerna exempelvis uttryckas som följer:

$$u(z) = 1 - e^{-kz} \quad k > 0$$

där k kallas för riskavertkonstant, ju större k desto mer riskavert beteende.

Exempel 12. Antag att den nuvarande förmögenheten är 0 så att den totala förmögenheten är lika med det vi tjänar på z .

Antag att \mathbf{c} är en slumpvariabel med väntevärde $\bar{\mathbf{c}}$ och kovariansmatris \mathbf{V} . Då är z en normalfördelad stokastisk variabel med väntevärde $\bar{z} = \bar{\mathbf{c}}^t \mathbf{x}$ och varians $\sigma^2 = \mathbf{x}^t \mathbf{V} \mathbf{x}$ och täthetsfunktionen är

$$\phi(z) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{z-\bar{z}}{\sigma}\right)^2}$$

Vi vill maximera väntevärdet av nyttan som ges av

$$\begin{aligned} \int_{-\infty}^{\infty} (1 - e^{-kz}) \phi(z) dz &= 1 - \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-kz} \frac{1}{\sigma} e^{-\frac{1}{2}\left(\frac{z-\bar{z}}{\sigma}\right)^2} dz = \\ &= 1 - e^{-k\bar{z} + \frac{1}{2}k^2\sigma^2} \end{aligned}$$

Nu ser vi att genom att maximera $k\bar{z} - \frac{1}{2}k^2x^tVx$ så maximerar vi nyttan. Alltså får vi följande kvadratsika programmeringsproblem

$$\text{Maximera } \mathbf{k}\bar{\mathbf{c}}^t\mathbf{x} - \frac{1}{2}\mathbf{k}^2\mathbf{x}^t\mathbf{V}\mathbf{x}$$

$$\text{Då } \begin{array}{l} \mathbf{A}\mathbf{x} \leq \mathbf{b} \\ \mathbf{x} \geq \mathbf{0} \end{array}$$

△

15.2 Stokastisk programmering

En typ utav storskalig optimering som är vanligen förekommande är *stokastisk programmering*, vilket också kan ses som ett två-steps stokastiskt optimal-kontrollproblem.

Först tar man en vektor $x \in \mathbf{R}^n$ från en mängd X som har vissa restriktioner och en slumpartad händelse inträffar, händelsen har m möjliga utfall, numrerade $1, \dots, m$. En annan vektor $y_i \in \mathbf{R}^m$ väljs, med vetskapen om utfallet i i , under villkoret

$$\mathbf{A}x + \mathbf{B}_iy_i = b_i, \quad y_i \geq 0, \quad i = 1, \dots, m$$

där matriserna \mathbf{A} och \mathbf{B} och vektorerna b_i är givna. Problemet är att minimera den förväntade kostnaden

$$f_0(x) + \sum_{i=1}^m \pi_i d_i^t y_i$$

där $f_0(x)$ är kostnaden för valet x , $d_i^t y_i$ är den kostnad som är knuten till händelsen i och det y_i man valt och π_i är den motsvarande sannolikheten.

Problemet kan nu skrivas som

$$\text{Minimera } f_0(x) + \sum_{i=1}^m \min_{B_i y_i = b_i - Ax, y_i \geq 0} \pi_i d_i^t y_i$$

$$\text{Då } x \in X$$

Vi kan nu ersätta minimeringsproblemet med dess duala problem:

$$\text{Minimera } f_0(x) + \sum_{i=1}^m \max_{B_i^t \lambda_i \leq \pi_i d_i} (b_i - Ax)^t \lambda_i$$

$$\text{Då } x \in X$$

16 Finansiella applikationer

Genom att simulera en marknad som en mängd stokastiska processer kan vi använda (och utveckla) det vi hittills lärt oss i denna uppsats till att säga något om det vi i inledningen sade oss vilja, nämligen ge matematiskt grundade råd för vad en tillgång bör kosta och vilken avkastning en investering bör ge i framtiden. Vi inleder kapitlet med att ”definera” en marknad och en tillgång och går därefter in på beräkningar givet våra definitioner.

16.1 Marknaden

Vi kommer att betrakta en marknad som en mängd av stokastiska processer som utvecklas med tiden.

Definition 16.1. En *marknad* är en $(n+1)$ -dimensionell Itôprocess $X(t) = (X_0(t), X_1(t), \dots, X_n(t))$ där $dX_0(t) = \rho(t)X_0(t)dt$ är en riskfri tillgång med $X_0(0) = 1$ och för $1 \leq i \leq n$:

$$dX_i(t) = \mu_i dt + \sigma_i dB(t), \quad X_i(0) = x_i$$

där $B(t)$ är en brownisk rörelse. Enklast är att se X_0 som ett bankkonto och X_1, \dots, X_n som aktier, men de kan naturligtvis även representera andra riskabla tillgångar.

Eftersom vi alltid vill diskontera till s.k. nuvärden så uttrycker vi från och med nu den säkra investeringen som:

$$dX_0(t) = \rho(t)X_0(t)$$

Eller om man så vill:

$$X_0(t) = \exp\left(\int_0^t \rho(s)ds\right)$$

Om vi vill *normalisera*¹² marknaden, dvs uttrycka priset på de riskabla tillgångarna i priset på den säkra tillgången, definerar vi

$$\xi(t) = X_0^{-1}(t) = \exp\left(-\int_0^t \rho(s)ds\right)$$

Och sätter sedan

$$\overline{X}_i(t) = \xi(t)X_i(t)$$

Då är

$$\overline{X}(t) = (1, \overline{X}_1(t), \dots, \overline{X}_n(t))$$

en normaliserad marknad.

¹²Detta innebär att $X_0(t) := 1 \forall t$

Definition 16.2. En *portfölj* är en $(n+1)$ -dimensionell process¹³

$$\theta(t) = (\theta_0(t), \theta_1(t), \dots, \theta_n(t))$$

En portfölj θ har vid varje tidssteg t värdet:

$$V_t^\theta(t, \omega) = \theta(t) \cdot X(t) = \sum_{i=0}^n \theta_i(t) X_i(t) \quad (16.1)$$

Definition 16.3. En portfölj $\theta(t)$ är *självfinansierande* om:

$$V^\theta(t) = V^\theta(0) + \int_0^t \theta(s) \cdot dX(s) \Leftrightarrow dV^\theta = \theta(t) \cdot dX(t)$$

Det vill säga vi lägger inte till eller tar ut några pengar ur vår investering. Värdeförändringen i vår portfölj med tiden är enbart beroende av förändringen i pris på de stokastiska X_i som däri ingår.

Anmärkning 16.4. För alla portföljer som vi skall studera kommer vi att kräva två saker, för det första att den är tillåtlig¹⁴, vilket innebär att vi ställer ett krav på vår portfölj som matematiskt kan uttryckas:

$$\exists K < \infty : V^\theta(t, \omega) \geq -K \quad \forall (t, \omega) \in [0, T] \times \Omega$$

Och för det andra att marknaden som portföljen hämtas från saknar arbitragemöjligheter, som i symboler skrivs:

$$\neg[(V^\theta(0) = 0) \wedge (V^\theta(T) \geq 0) \wedge (P(V^\theta(T) > 0))]$$

Där det första kriteriet tolkas som att vi ställer en maxgräns för hur mycket kredit som tolereras för en aktör på marknaden och det andra kriteriet kommer från antagandet att ingen fungerande marknad tillåter en aktör att riskfritt tjäna pengar utan någon initial personlig investering.

Exempel 13. Om vi kombinerar Definition 16.3 och (16.1) får vi följande uttryck:

$$\begin{aligned} V^\theta(t) &= \theta_0(t)X_0(t) + \sum_{i=1}^n \theta_i(t)X_i(t) = \\ &= \theta_0(0)X_0(0) + \int_0^t \theta_0(s)dX_0(s) + \int_0^t \sum_{i=1}^n \theta_i(s)dX_i(s) \Leftrightarrow \end{aligned}$$

¹³Processen är anpassad till den m -dimensionella σ -algebran som skapas av $B(t)$

¹⁴Från engelskans *admissible*

$$\begin{aligned} &\Leftrightarrow \theta_0(t)X_0(t) - \theta_0(0)X_0(0) = dV^\theta(t) = \\ &= \int_0^t \theta_0(s)dX_0(s) + \underbrace{\sum_{i=0}^n \left(\int_0^t \theta_i(s)dX_i(s) - \theta_i(t)X_i(t) \right)}_{=A(t)} \end{aligned}$$

Om vi nu låter $Y_0(t) = \theta_0(t)X_0(t)$ och $A(t)$ som ovan så får vi:

$$dY_0(t) = \rho(t)Y_0(t)dt + dA(t)$$

Som löses av:

$$\xi(t)Y_0(t) = \theta_0(t) = \theta_0(0) + \int_0^t \xi(s)dA(s)$$

Om vi använder Sats 5.5 på ovanstående får vi:

$$\begin{aligned} \theta_0(t) &= \theta_0(0) + \xi(t)A(t) - A(0) - \int_0^t A(s)d\xi(s) = \\ &= V(0) + \xi(t)A(t) - \int_0^t \rho(s)A(s)\xi(s)ds \end{aligned} \quad (16.2)$$

Om vi använder $\rho = 0$ ovan så får vi

$$\theta_0(t) = V(0) + \bar{A}(t)$$

Så vi kan om vi sammanfattar göra varje portfölj självfinansierande om vi givet $\theta_1, \dots, \theta_n$ väljer θ_0 enligt (16.2). \triangle

För att undvika att de marknader vi simulerar inte uppfyller Anmärkning 16.4, dvs tillåter arbitrage, vill vi ha en enkel metod för att kontrollera att så inte är fallet. På en marknad som inte uppfyller Anmärkning 16.4 är det nämligen möjligt att generera alla möjliga värden vid sluttiden oavsett vilken initial förmögenhet vi har (se [1] s. 251-253). Om så är fallet saknar varje beräkning mening i det att den inte hjälper oss tolka den reella finansiella världen. Därför behöver vi:

Sats 16.5. *Marknaden $X_t = (X_1(t, \omega), \dots, X_n(t, \omega))$ saknar arbitrage om och endast om det finns en Itôintegrerbar m -dimensionell process $u(t, \omega)$ sådan att*

$$\sigma(t, \omega)u(t, \omega) = \mu(t, \omega) - \rho(t, \omega)X(t, \omega) \quad (16.3)$$

och

$$E\left[\exp\left(\frac{1}{2} \int_0^t u^2(t, \omega)ds\right)\right] < \infty$$

Bevis. Vi visar först \Leftarrow
 Antag $\rho = 0$. Definera

$$dQ := \exp\left(-\int_0^t u dB(s) - \frac{1}{2} \int_0^t u^2 ds\right) dP \quad (16.4)$$

Då är Q ekvivalent med P , och därför finns en brownisk rörelse

$$\hat{B}(t) = \int_0^t u(s, \omega) ds + B(t) \quad (16.5)$$

Nu är

$$dX_i(t) = \mu_i dt + \sigma_i dB(t) = \sigma_i u dt + \sigma_i dB(t) = \sigma_i d\hat{B}(t)$$

$\hat{B}(t)$ är en martingal enligt Exempel 2 och därför är även $X_i(t)$ en Q -martingal.

Antag nu att $\theta(t)$ ger arbitrage på X_t och att $\bar{V}^\theta(0) = 0$. Då gäller

$$d\bar{V}^\theta = \xi dV^\theta + V^\theta d\xi = \xi \theta dX + \rho \xi V^\theta dX = \theta d\bar{X}$$

Alltså är \bar{V}^θ en undre begränsad martingal med avseende på Q . En sådan martingal är en supermartingal¹⁵, dvs

$$E[V^\theta(T)] \leq V^\theta(0)$$

Men detta är en motsägelse mot att θ ger arbitrage på $X_t \Leftrightarrow E[V^\theta(T)] > 0$, eftersom $V^\theta(0) = 0$.

För att visa \Rightarrow antar vi först att X_t saknar arbitrage och definierar sedan en mängd F till att innehålla $\omega \in \Omega$ sådana att (16.3) saknar lösning. Sen bestämmer vi ett θ_i som håller sig positiv och studerar

$$dV^\theta = V^\theta(t) - V^\theta(0) = \int \sum \theta_i dX_i \geq 0$$

Med $V^\theta(0) = 0$ och antagandet om inget arbitrage måste $P(F = \emptyset) \rightarrow 1$. \square

16.2 Fordringar

Nu är vi redo att ge oss in på den mest ekonomiskt tillämpbara delen i detta kapitel, vi skall nu studera fordringar och deras matematiskt mest intressanta specialfall, optioner.

Definition 16.6. Givet att det finns u, Q, \hat{B} som definierats i (16.3), (16.4) och (16.5) så definierar vi:

¹⁵Se [1] övning 7.12

- a) En *fordran* är en (undre begränsad) stokastisk variabel, härnäst efter benämnd $F(\omega)$
- b) En $F(\omega)$ som ovan är *nåbar* om det existerar en tillåtlig portfölj θ och ett reellt tal z s.a:

$$F(\omega) = V_z^\theta(T) = z + \int_0^T \theta(t) dX_t$$

och

$$\bar{V}_z^\theta(t) = z + \int_0^t \xi(s) \sum_{i=1}^n \theta_i(s) \sigma_i(s) d\hat{B}(s)$$

- c) En marknad är *komplett* om varje fordran är nåbar

Del **b)** kan förklaras som att F är nåbar om vi med initial förmögenhet z kan finna en portfölj som uppfyller

$$F(\omega) = V_z^\theta(T, \omega)$$

Om det finns en portfölj θ som uppfyller **b)** ovan kallar vi den, i brist på bra svensk motsvarighet, en portfölj som *hedgar*¹⁶ den givna fordran F .

Att kunna hedga portföljer är av avgörande värde för att kunna prissätta optioner. Vi skall nu visa två resultat om optioners pris, men vi inleder med en definition.

Definition 16.7. En *option* på fordran F med sluttid T är garantin att få summan F vid tiden $t = T$ (eller $t = \tau$, se nedan).

Vi skall skilja på *köpooptioner* och *säljoptioner*, som ger rätten att köpa respektive sälja den underliggande tillgången F till ett i förväg överenskommet pris, samt på *Europeiska* och *Amerikanska* optioner, där den europeiska varianten endast ger rätt att göra den avtalade transaktionen vid sluttiden T , medan den amerikanska ger innehavaren rätt att utföra transaktionen vid någon tidpunkt τ , $0 \leq \tau \leq T$.

Det första resultatet vi önskar visa är vad priset för en europeisk köpooption bör vara. Det finns ett mycket känt resultat¹⁷ för denna frågeställning och vi ämnar arbeta oss igenom det. I §18 kör vi några simuleringar över nedanstående formel med värden hämtade från den verkliga världen (OMX:s hemsida).

¹⁶Förslag på ord är t.ex. omgärdar, begränsar, ringar in. Portföljen så att säga stänger in värdet på fordran

¹⁷Black-Scholes formel

Sats 16.8. Givet en normaliserad marknad som innehåller två olika tillgångar X_0, X_1 som uppfyller

$$dX_0(t) = e^{rt}X_0, \quad r \in \mathfrak{R} \quad X_0(0) = 1$$

$$dX_1(t) = \mu X_1 dt + \sigma X_1 dB(t), \quad X_1(0) = x_1$$

Så gäller att priset, $p_E(F)$, på den europeiska köpoptionen $F(\omega)$ med lösenpris K , definierad som:

$$F(\omega) = (X_1(T) - K)^+, \quad K \in \mathfrak{R}$$

är:

$$p_E(F) = x_1 \mathbf{N}(d_1) - K e^{r(T-t)} \mathbf{N}(d_2)$$

Där \mathbf{N} är fördelningsfunktionen för en normalfördelad stokastisk variabel och d_1, d_2 ges av.

$$d_1 = \frac{\log(\frac{x_1}{K}) + (r + \frac{1}{2}\sigma^2)}{\sigma\sqrt{T-t}}$$

$$d_2 = \frac{\log(\frac{x_1}{K}) + (r - \frac{1}{2}\sigma^2)}{\sigma\sqrt{T-t}}$$

Bevis. (Obs: Detta bevis är långt och innehåller många teknikaliteter, men vi hoppas att läsaren ändå följer med i varje steg eftersom mycket av det vi hittills lärt oss kommer till användning här) Vi börjar med att anta att det eftersökta priset på vår option är $Y(t)$ vid tiden t . Vi vill också att detta pris skall vara en funktion $f : \mathfrak{R} \rightarrow \mathfrak{R}$:

$$Y(t) = f(X_1(t), t)$$

Idén är nu att vi bildar en portfölj som hedgar värdet på optionen. Vi önskar dock inte spela på marknaden. Därför bygger vi en portfölj som innehåller en köpoption Y , men också en kort¹⁸ position $-\theta_1$ i tillgången X_1 . Vi tar även en lång¹⁹ position θ_0 i X_0 . Vi får en portfölj som vid varje tid t har värdet:

$$V^\theta(t) = Y(t) - \theta_1(t)X_1(t) + \underbrace{\theta_0 X_0 e^{rt}}_{=e^{rt}\theta_0}$$

Vår portföljs värdeutveckling blir alltså (från och med nu undlåter vi att poängtera att vi är i tidssteget t):

$$dV = dY - \theta_1 dX_1 - X_1 d\theta_1 + \theta_0 r e^{rt} dt + e^{rt} d\theta_0$$

¹⁸Vi säljer/lånar ut ett antal av tillgången X_1

¹⁹Vi köper några enheter av den riskfria tillgången X_0

Vi kräver som vi nämnt ovan att portföljen är självfinansierande (dvs $X_1 d\theta_1 = e^{rt} d\theta_0$) och kan därför skriva om uttrycket som

$$dV = dY - \theta_1 dX_1 + \theta_0 r e^{rt} dt \quad (16.6)$$

För att finna ett mer hanterligt uttryck för dY använder vi Sats 5.3 och får:

$$\begin{aligned} dY &= \frac{\partial f}{\partial x} dX_1 + \frac{\partial f}{\partial t} dt + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} (dX_1)^2 = \\ &= \frac{\partial f}{\partial x} dX_1 + \frac{\partial f}{\partial t} dt + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} \sigma^2 X_1^2 dt \end{aligned}$$

Nu kan vi skriva (16.6) som:

$$dV = \left(\frac{\partial f}{\partial x} - \theta_1 \right) dX_1 + \left(\frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} \sigma^2 X_1^2 + \theta_0 r e^{rt} \right) dt \quad (16.7)$$

Nu kan vi låta portföljen vara helt riskfri under en kort tid om vi väljer θ_1 sådan att (16.7) tappar sin dX_1 -term. Vidare har vi antagandet om arbitragefri marknad som ger oss att under detta korta tidssteg som vi nu undersöker portföljen måste den ge samma avkastning som den riskfria tillgången X_0 , dvs $dV = rV dt$. Så nu kan vi ge (16.7) utseendet:

$$rV dt = \left(\frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} \sigma^2 X_1^2 + \theta_0 r e^{rt} \right) dt \quad (16.8)$$

och icke-arbitrage resonemanget använt på (16.6) ger att

$$\theta_0 r e^{rt} dt = r(V - Y + \theta_1 X_1)$$

Som med (16.8) ger (dividerat med dt):

$$\begin{aligned} rV &= \frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} \sigma^2 X_1^2 + r(V - Y + \theta_1 X_1) \Leftrightarrow \\ &\Leftrightarrow \frac{\partial f}{\partial t} = r f - r X_1 \frac{\partial f}{\partial x} - \frac{1}{2} \frac{\partial^2 f}{\partial x^2} \end{aligned} \quad (16.9)$$

Som är en partiell differentialekvation som om vi löser den ger oss den eftersökta formeln $f(X_1, t)$. Sedan tidigare har vi också villkoret $f(x, T) = \max(x - K, 0)$. För att lösa den ovanstående ekvationen sätter vi först:

$$x = K e^\psi, \quad t = T - \frac{\tau}{\frac{1}{2}\sigma^2}, \quad f = K v(\psi, \tau), \quad L_1 = \frac{r}{\frac{1}{2}\sigma^2}$$

Nu får (16.9) utseendet:

$$\frac{\partial v}{\partial \tau} = \frac{\partial^2 v}{\partial \psi^2} + (L_1 - 1) \frac{\partial v}{\partial \psi} - L_1 v \quad (16.10)$$

Och vårt tidigare villkor blir nu $v(\psi, 0) = \max(e^\psi - 1, 0)$. Vi vill ha denna ekvation på formen

$$\frac{\partial w}{\partial \tau} = \frac{\partial^2 w}{\partial \psi^2} \quad (16.11)$$

Eftersom det är en ekvation som har standardlösning. Om vi sätter:

$$v = \exp\left(-\frac{1}{2}(L_1 - 1)\psi - \frac{1}{4}(L_1 + 1)^2\tau\right)w(\psi, \tau)$$

Så kan läsaren undersöka att w nu uppfyller (16.11) när vi sätter in den i (16.10) och differentierar. Standardlösningen för ett sådant w som vi nu skapat är:

$$w(\psi, \tau) = \frac{1}{2\sqrt{\pi\tau}} \int_{-\infty}^{\infty} w(\alpha, 0) \exp\left(-\frac{(\psi - \alpha)^2}{4\tau}\right) d\alpha \quad (16.12)$$

Där

$$w(\psi, 0) = \max\left(\exp\left(\frac{1}{2}(L_1 + 1)\psi\right) - \exp\left(\frac{1}{2}(L_1 - 1)\psi\right), 0\right)$$

Vi gör variabelbytet:

$$\iota = \frac{\alpha - \psi}{\sqrt{2\tau}} \Rightarrow d\alpha = \sqrt{2\tau} d\iota$$

Och studerar (16.12) som nu har utseendet:

$$\begin{aligned} w(\psi, \tau) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} w(\psi + \iota\sqrt{2\tau}, 0) \exp\left(-\frac{1}{2}\iota^2\right) d\iota = \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\frac{\psi}{\sqrt{2\tau}}}^{\infty} \underbrace{\exp\left(\frac{1}{2}(L_1 + 1)(\psi + \iota\sqrt{2\tau})\right) e^{-\frac{1}{2}\iota^2}}_{=I_1} d\iota + I_2 \end{aligned}$$

Där I_2 är identisk med I_1 så när som på att termen $(L_1 + 1)$ blivit $(L_1 - 1)$. Att lösa I_1 och I_2 kräver flera steg och variabelbyten och vi nöjer oss med att konstatera att

$$I_1 = \frac{\exp\left(\frac{1}{2}(L_1 + 1)\psi + \frac{1}{4}(L_1 + 1)^2\tau\right)}{\sqrt{2\pi}} \int_{-\frac{\psi}{\sqrt{2\tau}} - \frac{1}{2}(L_1 + 1)\sqrt{2\tau}}^{\infty} e^{-\frac{\rho^2}{2}} d\rho$$

Och sen löser vi ut ψ och τ ur ursprungantagandena:

$$\psi = \log\left(\frac{x}{K}\right), \quad \tau = \frac{1}{2}\sigma^2(T-t), \quad f = Kv(\psi, \tau)$$

Som ger

$$f(x, t) = x\mathbf{N}(d_1) - Ke^{-r(T-t)}\mathbf{N}(d_2)$$

Med d_1 och d_2 som i Sats 16.8. □

Efter dessa långa beräkningar kastar vi oss direkt på ett besläktat resultat, vi ska undersöka vad det rätta priset på en amerikansk köpoption är.

Sats 16.9. *Givet Q som i (16.4) och att priset på en amerikansk köpoption beskrivs som*

$$p_A(F) = \sup_{\tau \leq T} E_Q[e^{-\rho\tau}(X_1(\tau) - K)^+]$$

på en marknad som är identisk marknaden given i Sats 16.8 så gäller att $p_A(F) = p_E(F)$.

Dvs priset för en amerikansk option är samma som det för en europeisk option och kan alltså beräknas med hjälp av Sats 16.8

Bevis. Vi vill visa att

$$\sup_{\tau \leq T} E_Q[e^{-\rho\tau}(X_1(\tau) - K)^+] = E_Q[e^{-\rho T}(X_1(T) - K)^+]$$

Dvs att:

$$E_Q[e^{-\rho t_1}(X_1(t_1) - K)^+] \leq E_Q[e^{-\rho t_2}(X_1(t_2) - K)^+], \quad \forall T \geq t_2 \geq t_1 \geq 0 \quad (16.13)$$

Vi inleder med att definera

$$Y(t) := (X_1(t) - K)$$

Och vill visa

$$E_Q[Y(t)] \leq E_Q[Y(t+h)], \quad h \geq 0 \quad (16.14)$$

Vi skriver om (16.4):

$$\begin{aligned} dQ &= \exp\left(-\int_0^t u dB(s) - \frac{1}{2}\int_0^t u^2 ds\right) dP \Leftrightarrow \\ &\Leftrightarrow dP = \exp\left(\frac{1}{2}\int_0^t u^2 ds + \int_0^t u dB(s)\right) dQ \end{aligned}$$

Och enligt def av väntevärden får vi:

$$E_P[Y(t)] = \int_0^t Y(s) dP \Rightarrow$$

$$\Rightarrow E_Q[Y(t)] = \int_0^t Y(s) \underbrace{\exp\left(\int_0^t u dB(s) + \frac{1}{2} \int_0^t u^2 ds\right)}_{=I(t)} dQ$$

Vi observerar att $I(t)$ är större än 0 för alla $t \geq 0$ och att $I(t+h) \geq I(t), \forall h \geq 0$, dvs vi drar slutsatsen att (16.14) gäller.

Nu vill vi visa att $Z(t)$ definierad nedan också är en submartingal med avseende på Q .

$$Z(t) := (X_1(t) - K)^+$$

Om nu $Y(t) > 0$ så är $Y(t) = Z(t)$ och om $Y(t) \leq 0$ så är $Z(t) = 0$. Eftersom $()^+$ är en konvex funktion så ger Jensens olikhet²⁰ för alla fall då $Y(t) > 0$

$$Z(t) = Y(t) \leq E[Y(t+h)] \leq E[Z(t+h)] \Rightarrow Z(t) \leq E_Q[Z(t+h)]$$

Så vi drar slutsatsen att även $Z(t)$ är en Q -submartingal.

Låt nu

$$f = (e^{-\rho t}(X_1(t) - K)^+)$$

Eftersom $Z(t)$ är en submartingal²¹ har vi alltså att $f(x) \leq E_x[f(X_\tau)]$.

Vi beräknar (med $t_2 > t_1$):

$$E_Q[f(X(t_2)) | X(t_1)] = E_{X(t_1)}[f(X_{t_2-t_1})] \geq f(X(t_1))$$

Antag nu att $t_1 \leq t_2$ är två tider i intervallet $[0, T]$, Enligt ovanstående resonemang gäller nu:

$$E_Q[e^{-\rho t_1}(X_1(t_1) - K)^+] \leq E_Q[e^{-\rho t_2}(X_1(t_2) - K)^+]$$

Och vi har visat att det matematiskt sett alltid är lönsammast att lösa in en amerikansk köpoption vid sluttiden T , och att dess pris därmed kan uttryckas som p_E i Sats 16.8 ovan. \square

²⁰ $\vartheta : \mathfrak{R} \rightarrow \mathfrak{R}$ konvex så gäller $\vartheta(E[X]) \leq E[\vartheta(X)]$

²¹Resultatet är en aning djupare än vi går in på här, det bygger på [1] Thm 7.1.2 och det omvända beviset för Lemma 10.1.3 i samma text

17 Dualitet och portföljoptimering

I det här kapitlet utgår vi från frågeställningen: Hur maximerar vi nyttan²² av en investering, eller, hur bygger vi en portfölj så att vi maximerar den förväntade avkastningen? Vi kommer att länka ihop de två hittills åtskilda delarna av denna uppsats i detta kapitel och visa att vi kan optimera en ekvation med stokastiska komponenter genom att använda oss av dualitet. Vi kommer använda oss av Definition 16.1 och Definition 16.3 även i detta kapitel men vi kommer modifiera Definition 16.2 en aning genom att låta:

$$\theta(t) = (\theta_1(t), \dots, \theta_n(t))$$

Vi kommer även i fortsättningen att låta den riskfria positionen vara θ_0 , men vi låter den inte ingå i $\theta(t)$ på grund av smidighetsskäl som snart kommer visa sig. Vi ger oss direkt på ett exempel.

Exempel 14. Vi antar att vi är en investerare med initial förmögenhet x kr. Genom att bygga en portfölj $\theta + \theta_0$ vill vi nu maximera vår förväntade nytta, dvs vi vill beräkna

$$\sup E[U(T)]$$

Vi antar att vår nyttofunktion U uppfyller de vanliga nationalekonomiska kraven²³. Om vi nu antar att vår nuvarande förmögenhet kan beskrivas som i §16, dvs som

$$V(t) = \theta(t) \cdot X(t) + \theta_0 r t$$

Kan vi se vår nuvarande förmögenhet som en process

$$dV(t) = rV(t)dt + \theta(\sigma dB(t) + (\mu - \mathbf{r})dt) \quad (17.1)$$

Där σ mäter de olika aktiernas varians och μ mäter väntevärdet av värdeökningen på de involverade aktierna. Det fetstilade \mathbf{r} :et är naturligtvis en $n \times 1$ vektor med enbart värdet r på varje position, så att dimensionerna stämmer²⁴. För bättre förståelse i notationen skriver vi från och med nu nyttofunktionen som

$$U(V(t))$$

Dvs vi ämnar maximera

$$\sup E[U(V(T))]$$

²²Detta är en ekonomisk term för att tjäna så mycket pengar som möjligt

²³Den är ökande, konkav och har avtagande derivata. Dessutom är den positiv för alla positiva värden. Funktionen \sqrt{x} är alltså en bra kandidat, då den uppfyller alla krav vi har, men även tidigare nämnda $1 - e^{-kx}$ är en tänkbar variant

²⁴För att dimensionerna ska stämma låter vi också σ vara $n \times n$, $B(t)$ vara en n -dimensionell Brownsk rörelse och μ är såklart en $n \times 1$ -vektor

för något fixt $T > 0$.

Lite eftertanke visar att

$$rV(t) + \theta(\mu - \mathbf{r}) = r\theta_0 + \mu\theta_1 + \mu\theta_2 + \cdots + \mu\theta_n$$

Ett resonemang som visar att portföljvärdet uppför sig som det "skall" när vi definierar det som ovan.

Vi ska nu försöka optimera funktionen genom Lagrangedualitet, för att göra det inför vi en process W som motsvarar Lagrangemultiplikatorn i den deterministiska analysen

$$dW(t) = W(t)(\alpha dt + \beta dB(t)) \quad (17.2)$$

Nu ser vi (17.1) som ett krav som V skall uppfylla och studerar

$$\int_0^T W(s)dV(s) \quad (17.3)$$

Om vi utvecklar (17.3) genom Sats 5.5 får vi²⁵

$$\begin{aligned} \int_0^T W_s dV_s &= [W_s V_s]_0^T - \int_0^T V_s dW_s - \int_0^T dV_s \cdot dW_s = \\ &= W_T V_T - W_0 V_0 - \int_0^T V_s W_s \alpha ds - \int_0^T V_s W_s \beta dB_s - \\ &- \int_0^T W_s (\alpha(rV_s + \theta(\mu - \mathbf{r})) (ds)^2 + (\alpha\theta\sigma + \beta rV_s + \beta\theta(\mu - \mathbf{r})) ds dB_s + \beta\theta\sigma (dB_s)^2) = \\ &= W_T V_T - W_0 V_0 - \int_0^T V_s W_s \alpha ds - \int_0^T V_s W_s \beta dB_s - \int_0^T W_s \theta \sigma \beta ds \quad (17.4) \end{aligned}$$

Där vi använt Sats 5.3 för att stryka de ingående $(dt)^2$ - och $dt dB_t$ -termerna.

Vidare kan vi också utveckla (17.3) enligt:

$$\int_0^T W_s dV_s = \int_0^T W_s (rV_s + \theta(\mu - \mathbf{r})) ds + \int_0^T W_s \theta \sigma dB_s \quad (17.5)$$

Om vi använder Sats 5.6 och tar väntevärdet av (17.4) får vi

$$E\left[\int_0^T W_s dV_s\right] = E\left[W_T V_T - W_0 V_0 - \int_0^T V_s W_s \alpha ds - \int_0^T W_s \theta \sigma \beta ds\right] \quad (17.6)$$

²⁵Hädanefter låter vi $V(t) = V_t$ för smidigare notation

Och pss ger (17.5):

$$E\left[\int_0^T W_s dV_s\right] = E\left[\int_0^T W_s(rV_s + \theta(\mu - \mathbf{r}))ds\right] \quad (17.7)$$

Eftersom (17.6) och (17.7) måste vara lika för varje godtagbart V gäller att:

$$\begin{aligned} \sup E[U(V_T)] &= \sup E[U(V_T) + W_0V_0 - W_TV_T + \\ &+ \int_0^T W_s(rV_s + \theta(\mu - \mathbf{r}))ds + \int_0^T V_sW_s\alpha ds + \int_0^T W_s\theta\sigma\beta ds] \end{aligned} \quad (17.8)$$

Nu har vi kokat ner det ursprungliga problemet till Lagrangefunktionen:

$$\begin{aligned} \Lambda(W) &= \sup_{V,\theta} E[U(V_T) + W_0V_0 - W_TV_T + \\ &+ \int_0^T W_s(V_s(r + \alpha))ds + \int_0^T W_s(\theta(\mu - \mathbf{r} + \sigma\beta))ds] \end{aligned} \quad (17.9)$$

För att vi skall få ett ändligt värde när vi söker supremum över positiva V, θ ser vi att nedanstående två krav måste vara uppfyllda:

$$\begin{cases} \alpha + r = 0 \\ \sigma\beta + \mu - \mathbf{r} = 0 \end{cases} \Rightarrow \begin{cases} \alpha = -r \\ \beta = \sigma^{-1}(\mathbf{r} - \mu) \end{cases}$$

Nu bildar vi den duala funktionen D till den ursprungliga U genom:

$$D(w) = \sup_v (U(v) - vw)$$

Och studerar (17.9) för D istället för U , vi får:

$$\begin{aligned} \Lambda(W) &= \sup_{V,\theta} E[D(W_T) + W_0V_0 + \\ &+ \int_0^T W_s(V_s(r + \alpha))ds + \int_0^T W_s(\theta(\mu - \mathbf{r} + \sigma\beta))ds] \end{aligned} \quad (17.10)$$

med samma bivillkor som ovan. Givet att dessa bivillkor håller får vi att det maximala värdet på (17.10) måste vara

$$\Lambda(W) = E[D(W_T) + W_0V_0]$$

Vi kan nu bilda det duala problemet

$$\inf_W \Lambda(W) = \inf_W E[D(W_T) + W_0V_0]$$

Vi drar oss till minnes (17.2) och inser att:

$$W_t - W_0 = W_t(\alpha dt + \beta dB_t) \Leftrightarrow W_t = W_0 \frac{1}{1 - \alpha dt - \beta dB_t}$$

Om vi multiplicerar funktionen med ξ så att alla priser uttrycks i den riskfria tillgången, dvs vi låter $\beta = 1$, och uttrycker W som en exponentiell funktion (en bra idé om vi vill skapa en positiv, växande funktion) så får det sökta W utseendet:

$$W_t = W_0 e^{-\int_0^t r ds} \xi \quad (17.11)$$

Som tillsammans med våra bivillkor ger oss en idé om hur vi maximerar vår förväntade nytta. Om vi nu använder oss av dualitetsteorierna från tidigare ansnitt så vet vi att vi kan lösa optimeringsproblemet, men vi överlåter det åt den intresserade läsaren. \triangle

Ett stokastiskt uttryck för hur en maximerad nytta beter sig är kanske inte alltid så informativt som man kan önska, vi skall därför också undersöka ett besläktat fall. Vi räknar även detta som ett exempel.

Exempel 15. Antag att vi har samma situation som i Exempel 14. Nu vill vi undersöka hur vi skall bete oss för att vara säkra på att vår förmögenhet vid sluttiden, V_T , är åtminstone lika stor som en i förväg bestämd variabel, M . Utvecklingen av vår förmögenhet är:

$$dV_t = \theta_0 V_t dt + \theta(\sigma dB_t + (\mu - r)dt)$$

Nu vill vi finna θ sådan att vi med så litet V_0 (dvs initial förmögenhet) som möjligt har att:

$$P(V_T \geq M) \rightarrow 1$$

För att göra detta bildar vi funktionen

$$u_0(v) = \begin{cases} 0, & v \geq 0 \\ -\infty, & v < 0 \end{cases}$$

Och försöker finna

$$\sup E[u_0(V_T - M)]$$

Detta är inte särskilt svårt, vi kan låna alla beräkningar och antaganden från Exempel 14 och vidare vet vi att dualen till u_0 är så enkel som $d_0 := -u_0$. Detta ger oss sambandet:

$$\sup E[u_0(V_T - M)] = \inf_W E[d_0(W_T) - W_T M + W_0 V_0] = \inf_{W \geq 0} E[W_0 V_0 - W_T M]$$

Vi söker nu minsta V_0 sådant att $W_0V_0 - W_TM \geq 0$ för $W \geq 0$

$$\begin{aligned} W_0V_0 - W_TM \geq 0 &\Leftrightarrow V_0 \geq \frac{W_TM}{W_0} \Rightarrow \\ \Rightarrow V_0 &\geq \sup_{W \geq 0} E\left[\frac{W_TM}{W_0}\right] \Rightarrow W_T = \frac{V_0W_0}{M} \end{aligned}$$

Vi vet från Exempel 14 att W_t uttrycks

$$dW_t = W_t(-rdt + \beta dB_t)$$

med

$$\sigma\beta + \mu - \mathbf{r} = 0 \tag{17.12}$$

Så vad vi är ute efter är helt enkelt, med gällande (17.12):

$$\sup E[MW_T]$$

Som om vi lånar svarsuttrycket från Exempel 14 även kan uttryckas

$$\sup E\left[W_0 e^{\int_0^T r ds} M\xi\right]$$

Detta får avsluta det för uppsatsen förenande kapitlet, men vi vill poängtera att vi endast visat en liten del av de resultat som kan nås när dualitet appliceras på en stokastisk miljö. \triangle

18 Datorsimulering

I detta avslutande kapitel ska vi med Matlabs hjälp simulera några processer som vi diskuterat tidigare i uppsatsen. Den som söker en anledning till varför detta är relevant behöver bara studera exempelvis (17.1) i kapitlet ovan. I nästan alla de exempel som vi i slutändan hamnat i då vi applicerat teorin i uppsatsen ingår en slumpmässig term. Eftersom en exakt beräkning därmed är omöjlig att utföra är det av yttersta vikt att kunna simulera de ingående processerna för att ha någon chans att kunna uttala sig om de förlopp vi önskar studera. Vi menar också att vi här visar på behovet av att förstå matematisk teori för att göra datorer meningsfulla. Omvänt så visar vi att med en god förståelse för teori kan datorer vara ett mycket användbart redskap för den som önskar studera stokastiska processer. Vi utvecklar en metod för att uppskatta de stokastiska integraler vi stött på samt kontrollerar hur bra uppskattningen i Sats 16.8 är för priset på en option. Alla koder för hur detta utförs ligger i §B.

18.1 Brownsk rörelse

Vi använder oss av samma metod för att simulera en Brownsk rörelse som vi använde för att definiera den i §3. Vi låter en funktion ta in tre värden, två som styr över vilket intervall vi önskar simulera rörelsen²⁶, och ett tredje som styr hur ofta användaren vill att funktionen skall "hoppa"²⁷. Eftersom vi vet att Brownsk rörelse har fördelningen $N(0,t)$ så begränsar vi även hoppen till att maximalt få den höjd som simuleringsintervallet har längd. I §C finns bilder över simuleringar för intervallet $[0,1]$, med 100, 1000 och 5000 hopp på vägen.

18.2 m-dimensionell Brownsk rörelse

När vi har kod för att simulera en Brownsk rörelse har vi inga som helst problem att simulera en m-dimensionell Brownsk rörelse, vi låter ett program ta in fyra värden, tre som i den 1-dimensionella Brownsk rörelsen och en fjärde som styr dimensionen, m. Sedan simulerar vi m stycken 1-dimensionella Brownska rörelser och låter dem vara kolonnerna i en matris. I §C finns bilder över simuleringar gjorda för 3-, 5-, och 10-dimensionella Brownska rörelser.

²⁶T.ex. om intervallet är $[a,b]$ så ger vi programmet värdena a och b

²⁷Detta värde kallar vi n

18.3 Den stokastiska integralen

Eftersom vi antagit att priset på riskabla tillgångar utvecklas bland annat som en integral över en Brownsk rörelse så är det för oss intressant att kunna skatta vad värdet av en sådan integral är. Vi kommer att lösa detta genom att använda definitionen av en stokastisk integral, se (4.2), och sedan låta Matlab sköta de långa summorna över små intervall. Detta program kommer att kräva tre delprogram som

- 1 Simulerar en Brownsk rörelse
- 2 Partitionerar intervallet i små steg
- 3 Uppskattar processen X :s värde i varje intervall

Eftersom vi redan har ett program som simulerar Brownsk rörelse återanvänder vi det för att få värden till delen ΔB_t i (4.2). Vi skriver ett sidoprogram som tar in samma parametrar som det som simulerar Brownsk rörelse, vars funktion är att dela upp intervallet $[a, b]$ i n lika stora delar. För stort n borde vi närma oss en bra uppskattning av fallet när $n \rightarrow \infty$. Det tredje hjälpprogrammet vi skriver har som uppgift att, givet en funktion, simulera värden för den givna funktionen över de n stycken intervall som vi skapat med den partitionerande funktionen. Givet dessa tre hjälpprogram har vi inga problem att numeriskt uppskatta den slumpmässiga integralen

$$\int_a^b f(\omega, t)dB(t)$$

Vi skriver vårt huvudprogram sådant att det som standard tar värdet $n = 1000$ (men användaren kan naturligtvis ändra detta för snabbare eller noggrannare beräkning). Huvudprogrammet tar sedan värdena a och b , dvs intervallet för integralen, och skickar de tre värdena i tur och ordning till programmet som simulerar en Brownsk rörelse, programmet som partitionerar intervallet och slutligen till programmet som uppskattar processens X :s värde. Sedan summerar huvudprogrammet de returnerade värdena enligt (4.2) och ger därmed en uppskattning av en stokastisk integral. I §A finns tabeller över skattade värden för den stokastiska integralen av funktionerna $f(t) = e^{B_t}$, $f(t) = B_t$ och $f(t) = \sigma t$, där σ är en konstant.

18.4 Itoprocesser

En Itöprocess har som vi beskrivit tidigare utseendet

$$dX(t) = f(\omega, t)dt + g(\omega, t)dB(t)$$

Nu när vi har ett program som uppskattar värdet av integralen över $dB(t)$ kan vi enkelt utöka detta program till att skatta värdet av en $It\bar{o}$ process. Det finns i Matlab en standardfunktion som numeriskt uppskattar värdet av en dt -integral och om vi samkör detta med vårt program som uppskattar en stokastisk integral får vi ut en numerisk skattning av $It\bar{o}$ processen X vid tiden t . I §A finns några tabeller över skattade $It\bar{o}$ processer. En naturlig fråga att ställa sig nu är: "Hur bra är vårt program på att simulera en verklig situation?". Eftersom vi när vi undersöker ett slumpmässigt fenomen av nöd tvingas acceptera att vi kommer få slumpmässiga svar är det inte helt trivialt att ge ett bra svar på den frågan, men om vi studerar tabellen i §A.2 ser vi att vi i alla fall lyckats åstadkomma en simulering som tar hänsyn till storleken på variansen σ . När σ minskar minskar också skillnaden i värdet på vår undersökta integral. Som kan ses i tabellen är skillnaden i svar då $\sigma = 0,7 := 0,0025$ medans när vi minskat ner till $\sigma = 0,3$ får en skillnad på endast 0,0009 mellan den största och minsta integralen.

18.5 Black-Scholes prisformel för optioner

Att numeriskt beräkna vad priset för en köpoption är är inte särskilt svårt när vi har Sats 16.8 och Matlab har en funktion för att numeriskt uppskatta värdet för en normalfördelning. Vi låter vårt program ta emot de fem värden, x_1, T, K, σ, r , som behövs för att kunna utföra beräkningarna i Sats 16.8 och låter programmet räkna ut svaret. En tabell över några skattade värden finns i §A. Vi har använt aktieprishistorik hämtat från OMX svenska hemsida. Vi har sedan skattat värdet på σ med funktionen `stdav.m` som återfinns i §B. För att göra en så bra skattning som möjligt vore det önskvärt att skatta standardavvikelse och väntevärde över ett lika långt tidsintervall som det skattade optionspriset löper över, fast "bakåt" i tiden istället. Eftersom vi har ett program som sköter beräkningarna åt oss och OMX:s hemsida ställer upp med prisinformation för långt tillbaka i tiden är det inga problem att utföra dessa beräkningar. De skattningar som återfinns i §A är dock gjorda för $T = 0,17$ år, då tyvärr tidsbrist uppkommit nära deadline och vi ansett det viktigare att slutföra kapitlen än att upprepa en skattning enligt beskriven procedur som den intresserade läsaren själv kan åstadkomma. Man bör också vara medveten om att de flesta matematiska modeller, även de vi använder, är uppbyggda så att de stipulerar att μ är positiv. När man genomför en skattning så som vi har gjort är chansen stor att den nuvarande finanskrisen medfört fallande aktiekurser under ett längre tidsintervall, dvs det μ som vi finner blir negativt. Eftersom μ inte är en faktor i Sats 16.8 ställer detta inte till några problem här, men den som önskar utföra beräkningar gällande utvecklingen av aktiekurser bör vara medveten om detta. Ett förslag

på lösning av problemet är exempelvis att använda ett mycket långt intervall bakåt i tiden.

Som kan ses i tabellerna i §A blir priserna negativa för stora värden på K och små värden på T . Detta faktum är ganska logiskt och talar för att vi har en bra uppskattning, en alltför "dålig" köption, dvs en med högt lösenpris, kan inte utgivaren räkna med att få betalt för utan måste tvärtom betala för att någon skall vilja teckna sig för. Dock kommer antagligen ingen sådan option någonsin att skrivas ut, eftersom den ger arbitrage för den som tecknar den.

Referenser

- [1] Øksendal, Bernt: *Stochastic Differential Equations*, 5th edition, Springer 2000
- [2] Sasane, Amol: *Mathematical Economics*, Stockholm 2008
- [3] Rogers, L.C.G.: *Duality in constrained optimal investment and consumption problems: a synthesis*, Cambridge 2001
- [4] Moon, Szepessy, Tempone, Zouaris: *Stochastic Differential Equations: Models and Numerics*, ? 2008
- [5] Bazaraa, Serali, Shetty: *Nonlinear Programming, Theory and Algorithms*, 2nd edition, Wiley 1979
- [6] Bertsekas, Nedić, Ozdaglar: *Convex Analysis and Optimization*, Athena Scientific 2003
- [7] Rudin, Walter: *Principles of Mathematical Analysis*, 3rd edition, McGraw-Hill 1976
- [8] Rockafeller, R T: *Convex Analysis*, Princeton 1970
- [9] Sontag, E D: *Mathematical Control Theory*, 2nd edition, Springer 1998
- [10] Rockafeller, Wadenski, *SIAM Journal on Control Optimization* 39, 2000

A Appendix1: Tabeller med Matlab-simuleringsvärden

A.1 Tabell över stokastisk integral

Den stokastiska integralen över några funktioner, skattade över intervallet $[0, 1]$ uppdelat i 1000 delintervall

$f(B_t, t) = e^{B_t}$	$f(B_t, t) = B_t$	$f(B_t, t) = 0.65t$
$fdB_t = 0.0307$	$fdB_t = -2.6850 \cdot 10^{-4}$	$fdB_t = -1.4834 \cdot 10^{-5}$
$fdB_t = -0.0352$	$fdB_t = -3.9687 \cdot 10^{-4}$	$fdB_t = 1.4310 \cdot 10^{-5}$
$fdB_t = 0.0126$	$fdB_t = -11.000 \cdot 10^{-4}$	$fdB_t = -1.8692 \cdot 10^{-6}$
$fdB_t = 0.0340$	$fdB_t = 3.8100 \cdot 10^{-4}$	$fdB_t = 3.5771 \cdot 10^{-7}$
$fdB_t = 0.0175$	$fdB_t = 2.2372 \cdot 10^{-4}$	$fdB_t = 4.8320 \cdot 10^{-6}$

A.2 Tabell över skattad Itoprocess

Skattning av Itōprocessen $dX(t) = \int_0^t \mu ds + \int_0^t \sigma dB(s)$, med $\mu = 0,05$ och för några olika värden på σ . Skattningen är gjord över intervallet $[0, 10]$ med 1000 delintervall.

$\mu = 0,05, \sigma = 0,70$	$\mu = 0,05, \sigma = 0,50$	$\mu = 0,05, \sigma = 0,30$
$dX(t) = 2.4990$	$dX(t) = 2.4998$	$dX(t) = 2.5004$
$dX(t) = 2.5003$	$dX(t) = 2.4994$	$dX(t) = 2.5002$
$dX(t) = 2.5015$	$dX(t) = 2.5010$	$dX(t) = 2.5000$
$dX(t) = 2.5006$	$dX(t) = 2.4996$	$dX(t) = 2.4997$
$dX(t) = 2.5003$	$dX(t) = 2.5001$	$dX(t) = 2.4995$

A.3 Tabeller över skattade optionspriser

För att kunna beräkna optionspriser med Sats 16.8 måste vi ha värdet på σ , vi skattar det värdet med hjälp av aktieprishistorik från OMX:s hemsida och formeln:

$$s = \sqrt{\frac{\sum x_i - \bar{x}}{N - 1}}$$

Vi kommer använda $r = 0,0375$ då riksbankens reporänta då detta skrivs ligger på 3,75 procent.

Nedanstående tabell²⁸ visar köptionspriser för HM:s aktie för några oli-

²⁸Alla tabeller har värden inklistrade från Matlab varför de följer den amerikanska notationen med punkter för decimaltal, medan resten av denna text använder sig av den svenska notationen med kommatecken, eftersom den är skriven på svenska

ka lösentider och lösenspriser. Det skattade värde som använts är $s = 0,1836$ och priset för en HM-aktie på dagen två månader innan detta skrevs var 313 SEK, varför vi använder det som vårt x_1 .

$T \downarrow K \rightarrow$	300	310	313	320	330
0.17	17.2426	11.8930	10.3367	6.7902	1.9237
0.25	19.8604	14.4887	12.9252	9.3607	4.4662
0.33	22.0531	16.6661	15.0975	11.5202	6.6054
0.42	24.2799	18.8802	17.3072	13.7189	8.7861
0.50	26.1073	20.6991	19.1231	15.5270	10.5816

Tabell över köptionspriser på Swedbanks aktie, med startvärdet 114,75 SEK och med skattad $s = 0,1666$

$T \downarrow K \rightarrow$	100	110	115	120	150
0.17	11.9502	6.0588	3.3825	0.8723	-11.2073
0.25	12.8317	6.9156	4.2235	1.6956	-10.5151
0.33	13.6019	7.6678	4.9636	2.4219	-9.8942
0.42	14.3829	8.4337	5.7188	3.1646	-9.2503
0.50	15.0231	9.0637	6.3410	3.7777	-8.7124

Tabell över köptionspriser på Fortums aktie, med startvärdet 32,25 EURO och med skattad $s = 0,0776$

$T \downarrow K \rightarrow$	27,5	30	32,5	35	40
0.17	3.5085	1.8633	0.3796	-0.9567	-3.2397
0.25	3.7079	2.0333	0.5172	-0.8537	-3.2088
0.33	3.8836	2.1854	0.6428	-0.7564	-3.1714
0.42	4.0630	2.3428	0.7753	-0.6510	-3.1240
0.50	4.2111	2.4742	0.8875	-0.5598	-3.0782

B Appendix 2: Matlab-koder

Brown.m

(simulerar en nxn-matris med brownsk rörelse)

```
function [A] = Brown(e,f,n)
```

```
B = [0];
```

```
a=1;
```

```
b=0;
```

```
c=0;
```

```
t = (f-e)/n;
```

```
for a = 1:n
```

```
  b = rand;
```

```
  if b > 0.50001
```

```
    c=rand*t;
```

```
  else
```

```
    c=-rand*t;
```

```
  end
```

```
  B(a+1)= c + B(a);
```

```
end
```

```
A=B;
```

bsch.m

(BS call option simulation)

```
function [y] = bsch(S,T,K,r,sig);
```

```
normal = inline('(1+erf(x/sqrt(2)))/2', 'x');
```

```
d1 = (log(S/K) + (r+0.5*sig*sig)*T/sig/sqrt(T));
```

```
d2 = (log(S/K) + (r-0.5*sig*sig)*T/sig/sqrt(T));
```

```
y = S*normal(d1)- K*exp(-r*T)*normal(d2);
```

Ito.m

```
function [x] = Ito(a,b)
```

```
y = quad(@dtfun,a,b);
```

```
z = StokInt(a,b);
```

```
x = y+z;
```

NBrown.m

```
(m-dimensionell brownian motion)
function [Z] = NBrown(a,b,n,m)
((a,b) intervall, n antal iterationer, m dimension)
X=zeros(n,m);
c=1;
d=0;
for d=0:n-1
Y = Brown(a,b,n);
for c = (d*m+1):((d+1)*m)
X(c) = Y(c-d*m);
end
Y=[0];
end
Z = X;
```

Part.m

```
(en partition av ett intervall)
function [T] = Part(a,b,n)
z = 1; (för att räkna upp)
n1 = n; (ändra för kortare/längre intervall)
c = 0;
d = b-a;
N = [];
t = d/(n1*n1);
for z = 1:(n1)
N(z+1) = z*t + a;
end
T = N; (N är nxn-matris)
```

dBfun.m

```
(funktionen  $f(X,t) = 0.5t$ )
function [g] = dBfun(a, b, n)
Pmatris = Part(a,b,n); (en partitionering)
Hmatris = []; (hjälpmatris m värden)
Bmatris = Brown(a,b,n);
v = 1; (något att summera över)
```

```

for v = 1:n
Hmatris(v) = 0.5*Pmatris(v);
end
g = Hmatris;

```

dtfun.m

(här läggs funktionen som beror av dt)

```

function [f] = dtfun(t)
r = 0.05;
f = r*t; (f = rt)

```

stdav.m

```

function [s] = stdav
(för fortums aktie)
b = vv;
B = [];
a=0;
A = xlsread('fortun.xls');
for a = 1:length(A)
B(a) = (A(a) - b)*(A(a) - b);
end
c = sum(B);
d = sqrt(c/(length(A)-1));
s = d/B(1); (för att få i procent av priset)

```

StokInt.m

(uppskattning av en stokastisk integral)

```

function [Int] = StokInt(a,b)
n=1000; ( kan ökas för finare partitionering)
Sum = 0;
DelSum = 0;
BR = Brown(a,b,n);
F = dBfun(a,b,n);
w=1; (något att summera över)
for w =1:n
DelSum = F(w)*(BR(w+1) - BR(w));

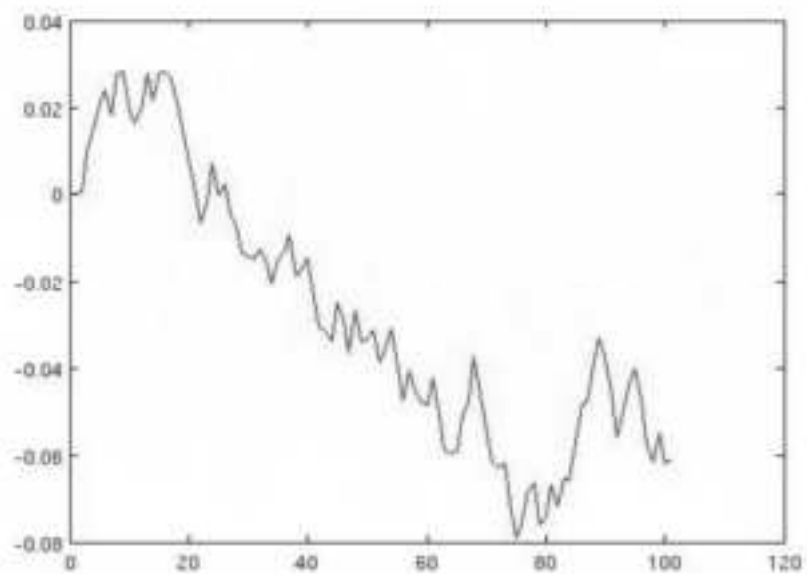
```

```
Sum = Sum + DelSum;  
end  
Int = Sum;
```

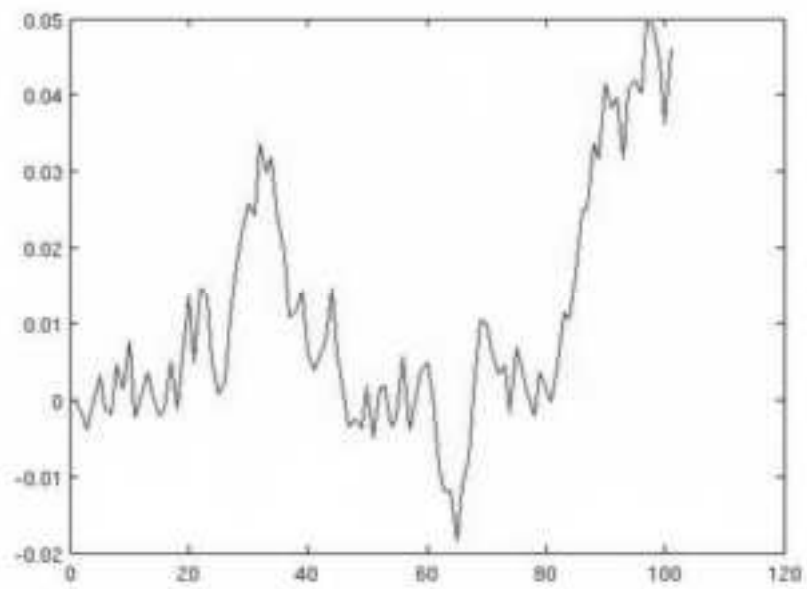
vv.m

```
function [m] = vv  
A = xlsread('fortun.xls');  
m = sum(A)/length(A);
```

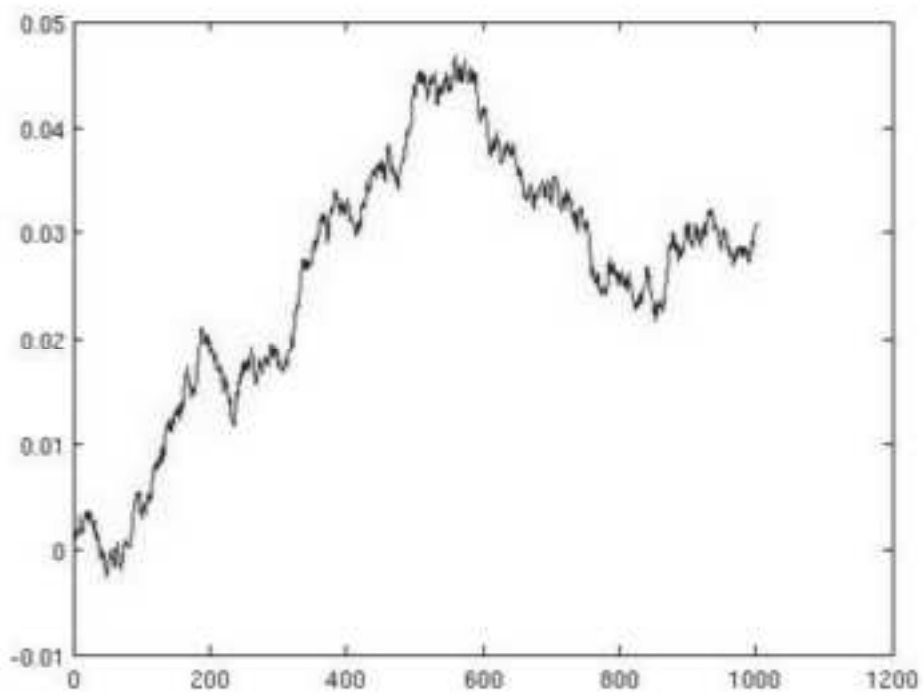
C Appendix3: Bilder



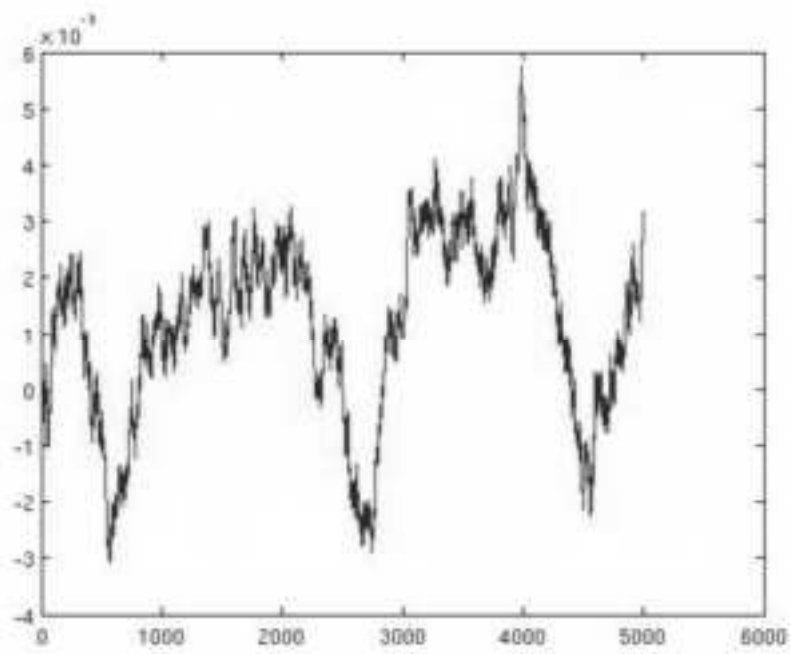
Figur 3: Brownsk rörelse med 100 delintervall



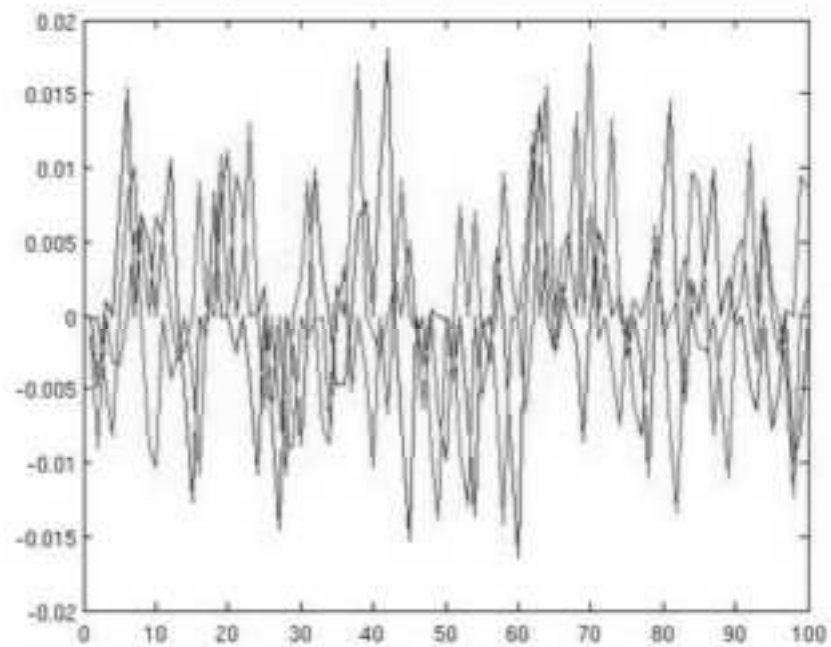
Figur 4: Brownsk rörelse med 100 delintervall



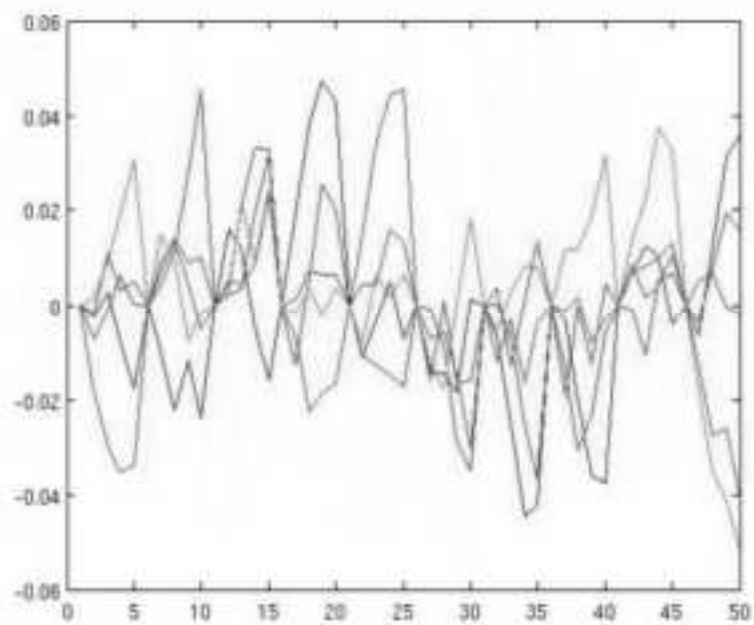
Figur 5: Brownsk rörelse med 1000 delintervall



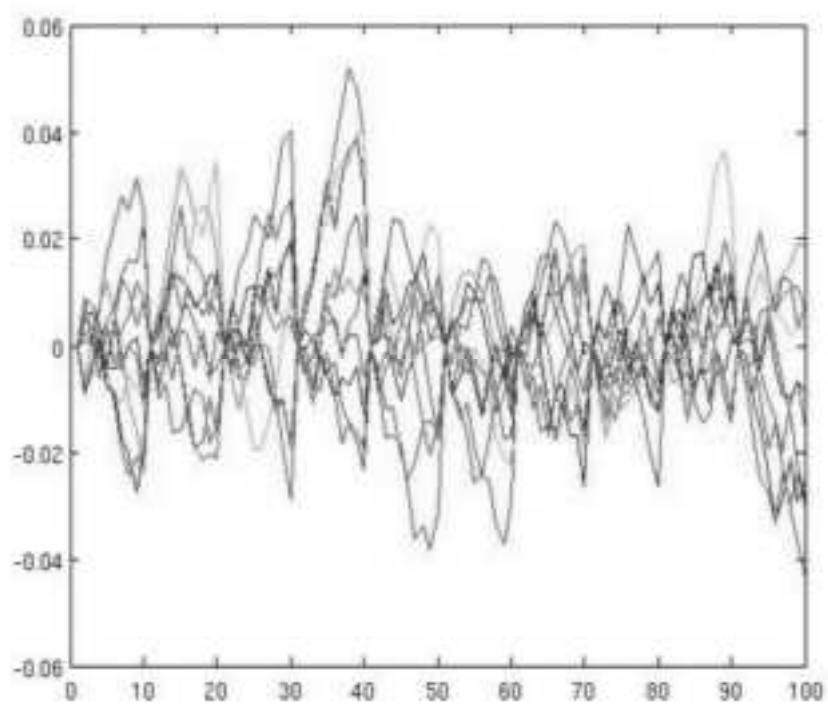
Figur 6: Brownsk rörelse med 5000 delintervall



Figur 7: en 3-dimensionell Brownsk rörelse med 100 delsteg



Figur 8: en 5-dimensionell Brownsk rörelse med 50 delsteg



Figur 9: en 10-dimensionell Brownsk rörelse med 100 delsteg