Tidsdiskreta system

Bengt Carlsson
Systemteknik
IT-institutionen Uppsala universitet

20 januari 2010

Sammanfattning

Detta material ger en introduktion till tidsdiskreta system och stokastiska processer

1 Tidskontinuerliga system

Vi börjar med en kort repetition från reglerteknikkursen när det gäller tidskontinuerliga system. Vi utgår från den allmänna linjära (tidsinvarianta) differentialekvationen

$$\frac{d^n y(t)}{dt^n} + a_1 \frac{d^{n-1} y(t)}{dt^{n-1}} + \ \ldots \ + a_n y(t) = b_0 \frac{d^m u(t)}{dt^m} + b_1 \frac{d^{m-1} u(t)}{dt^{m-1}} + \ \ldots + b_m u(t) \quad (1)$$

där $m \leq n$ Om systemet är i vila vid $t = 0$ (y och u och alla deras derivator har begynnelsevärdet 0) blir Laplacetransformen av (1)

$$(s^n + a_1 s^{n-1} + \ \ldots + a_n) Y(s) = (b_0 s^m + b_1 s^{m-1} + \ \ldots \ + b_m) U(s) \qquad (2)$$

där $Y(s)$ och $U(s)$ är Laplacetransformerna av signalerna $y(t)$ och $u(t)$. Systemets överföringsfunktion $G(s)$ fås genom att bilda kvoten mellan $Y(s)$ och $U(s)$ som

$$G(s) = \frac{Y(s)}{U(s)} = \frac{b_0 s^m + b_1 s^{m-1} + \ \ldots \ + b_m}{s^n + a_1 s^{n-1} + \ \ldots + a_n} \qquad (3)$$

Från transformteorin vi att en multiplikation i Laplacedomänen representerar en faltning i tidsdomänen. För sambandet

$$Y(s) = G(s) U(s) \qquad (4)$$

fås

$$y(t) = \int_0^t g(\tau) u(t - \tau) d\tau \qquad (5)$$

där $g(t) = \mathcal{L}^{-1}(G(s))$ kallas *viktfunktion* eller *impulssvaret*. Notera att om insignalen är en Diracpuls $u(t) = \delta(t)$, fås

$$y(t) = \int_0^t g(\tau)\delta(t-\tau)d\tau = g(t) \tag{6}$$

Systemets *poler* definieras som nollställena till överföringsfunktionens nämnarpolynom, det vill säga rötterna till

$$s^n + a_1 s^{n-1} + \ \ldots + a_n = 0$$

som också kallas den karakteristiska ekvationen. Vi har följande viktiga stabilitetsresultat:

Ett system är insignal-utsignalstabilt och asymptotiskt stabilt om alla poler har strikt negativ realdel.

Ett par viktiga räkneregler:

- *Slutvärdesteoremet.* För en signal $y(t)$ gäller att

  $$\lim_{t\to\infty} y(t) = \lim_{s\to 0} sY(s)$$

  om gränsvärdet $y(\infty)$ existerar, vilket gäller då $y(t)$ konvergerar mot ett konstant värde.

- *Fördröjningssatsen.* Antag att vi har ett system med en ren tidsfördröjning, $y(t) = u(t-T_d)$, det vill säga T_d är den tid som systemet fördröjer inverkan av insignalen. Då gäller att

  $$Y(s) = e^{-sT_d}U(s)$$

  alltså är Laplacetransformen för en ren tidsfördröjning T_d lika med e^{-sT_d}.

Om insignalen till ett linjärt system $G(s)$ är en sinussignal $u(t) = A\sin\omega t$, så blir även systemets utsignal (efter lång tid) en sinussignal enligt

$$\boxed{y(t) = A|G(i\omega)|\sin(\omega t + \arg(G(i\omega)))} \tag{7}$$

det vill säga utsignalens amplitud är förstärkt med en faktor $|G(i\omega)|$ och utsignalen är fasförskjuten $\arg(G(i\omega))$ radianer i förhållande till insignalen. Eftersom funktionen $G(i\omega)$ för en given frekvens ω bestämmer utsignalens beteende, kallas denna funktion för *frekvensfunktion*. Ett vanligt sätt att grafiskt presentera frekvensfunktionen $G(i\omega)$ är i ett så kallat Bodediagram. I ett Bodediagram plottas beloppskurvan $|G(i\omega)|$ (eller snarare 10-logaritmen av beloppskurvan) och argumentet $\arg(G(i\omega))$ mot logaritmen av vinkelfrekvensen

2 Tidsdiskreta system

Med ett tidsdiskret system menar vi ett system där utsignalen vid en viss tidpunkt beror av nuvarande och äldre värden på insignal och störning från vissa diskreta tidpunkter. Lite mer formellt kan man säga att en tidsdiskret signal är en uppräkningsbar sekvens. Ett tidsdiskret system är ett system som verkar på en tidsdiskret insignal och producerar en tidsdiskret utsignal. Ofta används tidsdiskreta modeller (differensekvationer) som approximationer av tidskontinuerliga system (jämför t ex Eulers metod inom numerisk analys). Notera att för att kunna simulera en modell av ett dynamiskt system i en dator måste modellen ekvationer lösas i tidsdiskreta tidpunkter.

Om man vill skatta en modell direkt från uppmätta (samplade) data så att man har tillgång till data i diskret tid med vissa tidsintervall mellan mätningarna. Det är därför naturligt att då använda en tidsdiskret modell. Oftast utförs samplingen med jämn hastighet, dvs (vinkel) samplingsfrekvensen ω_s [rad/s] är konstant (samplingsfrekvensen f_s [Hz]). Samplingstiden ges av $T_s = 1/fs = 2\pi/\omega_s$.

2.1 Överföringsfunktionen för tidsdiskreta system

Ett allmänt kausalt tidsdiskret LTI (linjärt tidsinvariant) system kan skrivas som

$$y(k) = \sum_{n=0}^{\infty} h(n)u(k-n); \quad k = 0, 1, 2... \tag{8}$$

Sekvensen $h(n)$, $n = 0, 1, 2, , ,$ anger hur olika värden på insignalen viktas för att generera utsignalen. Vanliga namn på $h(n)$ är impulssvar, pulssvar och viktsfunktion.

Att systemet är kausalt betyder att utsignalen inte beror av framtida värden på insignalen, dvs $y(k)$ beror inte på $u(s)$ där $s > k$. I framställningen har tidsintervallet mellan värdena normerats till en tidsenhet.

OBS, I läroboken Modellbygge och simulering används t (även) för att beteckna tidsdiskreta sekvenser, dvs man skriver $y(t)$ där $t = 0, 1, 2...$

Systembeskrivningen (8) är ofta användbar i teoretiska härledningar. Ett mera naturligt sätt att beskriva ett linjärt samband mellan insignal- och utsignalsekvens i diskret tid är att ange en differensekvation.

$$y(k)+a_1y(k-1)+\ldots+a_ny(k-n) = b_0u(k)+b_1u(k-1)+\ \ldots+b_nu(k-n) \tag{9}$$

eller om man skiftar båda leden n tidssteg framåt

$$y(k+n)+a_1y(k+n-1)+\ldots+a_ny(k) = b_0u(k+n)+b_1u(k+n-1)+\ \ldots+b_nu(k)$$

Om vi som tidigare förutsätter att systemet är i vila och z-transformerar får vi uttrycket

$$(z^n + a_1 z^{n-1} + \ldots + a_n)Y(z) = (b_0 z^n + b_1 z^{n-1} + \ \ldots + b_n)U(z)$$

och överföringsfunktionen $H(z)$ kan bildas som kvoten mellan $Y(z)$ och $U(z)$ enligt

$$H(z) = \frac{Y(z)}{U(z)} = \frac{b_0 z^n + b_1 z^{n-1} + \ \ldots + b_n}{z^n + a_1 z^{n-1} + \ldots + a_n} \tag{10}$$

och det gäller att

$$Y(z) = H(z)U(z) = \frac{b_0 z^n + b_1 z^{n-1} + \ \ldots + b_n}{z^n + a_1 z^{n-1} + \ldots + a_n} U(z) \tag{11}$$

Notera likheten med överföringsfunktioner för kontinuerliga system ($G(s)$). Det är dock betydligt enklare att beskriva ett system med tidsfördröjning. Antag att vi har en tidsfördröjning i systemet på p steg. Det är då bara att sätt $b_o = b_1 = ...b_p = 0$ i (9) alternativt $h(0) = h(1) = ...h(p) = 0$ i (8).

2.2 Användning av förskjutningsoperatorn

En alternativ representation av tidsdiskret system kan fås med hjälp av förskjutningsoperatorn(kallas också skiftoperatorn) q som för en allmän tidsdiskret signal $x(k)$ definieras av

$$qx(k) = x(k+1)$$

och mera allmänt

$$q^n x(k) = x(k+n)$$

Bakåtskiftoperatorn definieras analogt

$$q^{-1}x(k) = x(k-1)$$

Med hjälp av q-operatorn kan (9) skrivas

$$(1 + a_1 q^{-1} + \ldots + a_n q^{-n})y(k) = (b_0 + b_1 q^{-1} + \ \ldots + b_n u q^{-n})u(k)$$

Vi kan nu definiera *överföringsoperatorn*

$$H(q) = \frac{b_0 + b_1 q^{-1} + \ \ldots + b_n q^{-n}}{1 + a_1 q^{-1} + \ldots + a_n q^{-n}} = \frac{b_0 q^n + b_1 q^{n-1} + \ \ldots + b_n}{q^n + a_1 q^{n-1} + \ldots + a_n} \tag{12}$$

Överföringsoperatorn kallas ofta lite slarvigt för överföringsfunktionen. I räkningar fyller q-formalismen samma funktion som z-transformen och det är snarare en smaksak vilket representation som används. En liten fördel med q-operatorn är att insignal-utsignal sambandet ges i tidsplanet dvs

$$y(k) = H(q)u(k)$$

I kursen Empirisk modellering kommer vi i princip bara att använda q-formalismen

2.3 Ett exempel

Antag föjande sambandet mellan y och u

$$y(k) = y(k-1) + u(k-2)$$

vilket om vi skiftar ekvationen två tidssteg framåt kan skrivas som

$$y(k+2) - y(k+1) = u(k)$$

Med hjälp av q-operatorn fås

$$(q^2 - q)y(k) = u(k) \tag{13}$$

Vi kan därmed skriva sambandet med överföringsfunktion (överföringsoperatorn) enligt

$$y(k) = H(q)u(k) = \frac{1}{q^2 - q}u(k) \tag{14}$$

2.4 Statisk förstärkning

Analogt med det tidskontinuerliga fallet antar vi att insignalen är en konstant sekvens med värdet u_0 och utsignalen svänger in mot en konstant nivå $y(\infty) = \lim_{t\to\infty} y(k)$. Den statiska förstärkningen definieras

$$K = \frac{y(\infty)}{u_0} \tag{15}$$

Genom att utnyttja egenskaper för Z-transformen (detaljer utlämnas) fås

$$y(\infty) = \lim_{k\to\infty} y(k) = \lim_{z\to 1} (z-1)Y(z) = \lim_{z\to 1} (z-1)H(z)\frac{u_0 z}{z-1} = H(1)u_0 \tag{16}$$

och den statiska förstärkningen för detta system blir därför

$$K = \frac{y(\infty)}{u_0} = H(1) \tag{17}$$

Tolkningen av begreppet statisk förstärkning är densamma som i kontinuerlig tid, det vill säga hur mycket en konstant insignal förstärks av systemet då alla transienter avklingat.

2.5 Poler och stabilitet för tidsdiskreta LTI-system

Systemets *poler* definieras analogt med det tidskontinuerliga fallet som nollställena till överföringsfunktionens nämnarpolynom, det vill säga rötterna till

$$q^n + a_1 q^{n-1} + \ \ldots + a_n = 0$$

Vi skall nu ge två stabilitetsvillkor för tidsdiskreta LTI-system:

- Ett system är insignal-utsignalstabilt om alla systemets poler (rötter till den karakteristiska ekvationen) λ_i ligger strikt innanför enhetscirkeln i det komplexa talplanet (det vill säga att alla poler har belopp mindre än 1).

- Ett system är insignal-utsignalstabilt om viktfunktionen $h(k)$ uppfyller

$$\sum_{n=0}^{\infty} |h(n)| < \infty \tag{18}$$

2.6 Ett exempel på samband mellan polläge och respons i tidsplanet

Betrakta ett första ordningens tidsdiskret system som beskrivs av differensekvationen

$$y(k+1) = \gamma y(k) + (1-\gamma)u(k) \tag{19}$$

vilket med hjälp av överföringsfunktion kan skrivas som

$$y(k) = H(q)u(k) = \frac{1-\gamma}{q-\gamma}u(k)$$

och systemet har alltså en pol i $q = \gamma$.

Om vi vill använda ett enhetssteg som insignal så gäller att $u(k) = 1$, $k \geq 0$, och differensekvationen (19) reduceras till

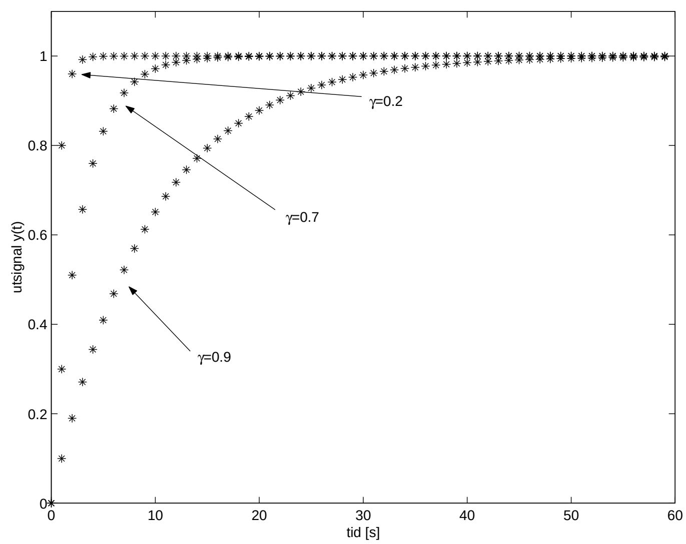
$$y(k+1) = \gamma y(k) + (1-\gamma), \quad k \geq 0 \tag{20}$$

Om vi vidare antar att $y(0) = 0$ kan via rekursion stegsvaret skrivas som

$$y(k) = (1-\gamma)\sum_{n=0}^{k-1} \gamma^n \tag{21}$$

och systemet är alltså stabilt om $\sum_{n=0}^{k-1} \gamma^n$ konvergerar mot något ändligt värde då $k \to \infty$, vilket gäller om $|\gamma| < 1$. Det värde summan konvergerar mot är då $\frac{1}{1-\gamma}$. Här har vi alltså för detta exempel verifierat stabilitetsdefinitionen.

Stegsvaret för systemet för några olika värden på γ visas i figur 1. Man ser att stegsvaret är snabbare ju närmare origo polen ligger.



Figur 1: Stegsvar för det tidsdiskreta systemet $\frac{1-\gamma}{q-\gamma}$ för några värden på γ

3 Tillståndsbeskrivning

Tillståndsbeskrivningar för tidsdiskreta LTI-system följer samma principer som för kontinuerliga system, med den skillnaden att man här skriver om en högre ordningens *differensekvation* som ett system av första ordningens differensekvationer. En tidsdiskret tillståndsbeskrivning har därför formen

$$\begin{aligned} x(k+1) &= Fx(k) + Gu(k) \\ y(k) &= H_d x(k) \end{aligned} \tag{22}$$

Polerna ges av egenvärdena till matrisen F. Överföringsfunktionen ges av

$$H(q) = H_d(qI - F)^{-1}G$$

4 Frekvensbeskrivning

Vi ska nu studera vad som händer med utsignalen till systemet (8) då insignalen är en tidsdiskret sinussignal (för enkelhets skull sätter vi amplituden till ett):

$$u(k) = \sin(\omega k)$$

I analysen är det bekvämt att skriva sinussignalen som $u(k) = \mathbf{Im}\ e^{i\omega k}$. Systembeskrivningen (8) ger då

$$
\begin{aligned}
y(k) &= \sum_{n=0}^{\infty} h(n)\mathbf{Im}\ e^{i\omega(k-n)} = \mathbf{Im}\ e^{i\omega k}\sum_{n=0}^{\infty} h(n)e^{-i\omega n} = \mathbf{Im}\ e^{i\omega k}H(e^{i\omega}) \\
&= \mathbf{Im}\ e^{i\omega k}|H(e^{i\omega})|e^{i\arg(H(e^{i\omega}))} = |H(e^{i\omega})|\sin(\omega k + \arg(H(e^{i\omega}))) \qquad (23)
\end{aligned}
$$

Det tidsdiskreta systemets frekvensfunktion erhålls då q ersätts med $e^{i\omega}$ i $H(q)$ Vi har samma tolkning som i det tidskontinuerliga fallet det vill säga att utsignalens amplitud är förstärkt med en faktor $|H(e^{i\omega})|$ och fasförskjuten $\arg(H(e^{i\omega}))$ radianer i förhållande till insignalen. På grund av symmetriskäl behöver man endast studera $H(e^{i\omega})$ för $0 \leq \omega \geq \pi$.

Vi har i analysen av de tidsdiskreta systemen antagit att tiden mellan två tidsdiskreta värden (t ex $y(k)$ och $y(k+1)$) är (normerat till) en tidsenhet. Om vi istället har en "verklig" tid mellan två närliggande värden på T_s $[s]$ erhålls rätt frekvensaxel om $H(e^{i\omega T_s})$ beräknas för $0 \leq \omega \geq \pi/h$.

4.1 Digitala filter

En mycket vanlig exempel på tidsdiskreta system är de algoritmer som används för digital signalbehandling. Ett typiskt exempel är man har en tidsdiskret signal där man vill framhäva och/eller undertrycka vissa egenskaper. T ex kan signalen innehålla slumpmässiga störningar som man vill reducera. Man talar ofta om digitala filter, som helt enkelt är ett tidsdiskret system som används för att behandla en tidsdiskret signal.

Ett exempel på ett digitalt filter är att beräkna ett glidande medelvärde baserat på N stycken mätningar. Om vi betecknar medelvärdet vid tiden k med y(k) kan vi skriva det tidsdiskreta systemet (digitala filtret) som

$$y(k) = \frac{1}{N}\sum_{p=0}^{p=N-1} u(k-p) \qquad (24)$$

där u betecknar den tidsdiskreta signal som filtreras. Ju större N väljs ju effektivare kommer slumpmässiga störningar att filtreras. Dock minskar förmågan att snabbt följa "verkliga" förändringar i signalen u då N ökar.

Ett mera allmänt (linjärt och tidsinvariant) digitalt filter kan skrivas

$$\sum_{m=0}^{M} a_m y(k-m) = \sum_{p=0}^{p=N} b_k u(k-p) \qquad (25)$$

Uppgiften är nu att designa koefficienterna $\{a_m\}_{m=0}^M$, $\{b_p\}_{p=0}^N$ så att filtret får lämpliga egenskaper.

I många fall är det naturligt att beskriva önskade egenskaper hos ett (digitalt) filter i frekvensplanet. Några typiska exempel:

- Den intressanta signalen har obetydlig energi över frekvensen ω_1 [rad/s]. Störningarna och brus har högre frekvenser.

- Mätningarna störs av en signal med frekvensen ω_2 [rad/s] (det kan t ex vara störningar från nätfrekvensen).

- För att kontrollera att inga resonanser uppkommer behöver frekvensen frekvensen ω_3 [rad/s] övervakas.

Ovanstående exempel leder till formuleringar där man vill att vissa frekvenser i en signal ska passera genom filtret medan andra ska stoppas. Det finns fyra klassiska typer av frekvensselektiva filter:

1. Lågpassfilter. Låga frekvenser i signalen ska passera och höga frekvenskomponenter ska ska spärras. Ett idealt lågpassfilter har förstärkning 1 för alla frekvenser upp till en viss gränsfrekvens ω_g och förstärkning noll för frekvenser över ω_g. Ett sådant filter kan inte realiseras med ett kausalt filter utan en approximation måste göras.

2. Högpassfilter. Höga frekvenser i signalen ska passera och låga frekvenser ska spärras.

3. Bandpassfilter. Signalens frekvenser inom ett visst frekvensområde ska passera medan övriga ska spärras.

4. Bandspärrfilter. Frekvenser inom ett visst frekvensområde ska spärras.

För att designa digitala filter finns bra datorstöd. En standardmetod i Matlab är funktionen *butter* som tar fram filterkoefficienter för ett s.k. ”Butterworthfilter”.

5 Tidsdiskreta stokastiska processer

Det är mycket vanligt att en tidsserie har en slumpmässig variation (eller inte kan beskrivas på något annat sätt). Detta kan t ex gälla naturföreteelser (vädervariatioer, havsvågor), tekniska system (störningar i elnätet), och ekonomiska system (börskursens utveckling).

För att beskrivning slumpmässiga förlopp hos en signal används normalt tidsdiskreta modeller (den matematiska beskrivningen blir väsentligen mycket enklare jämfört med om tidskontinuerliga modeller används). Ett standardredskap är att för att beskriva ett slumpmässigt förlopp är *stokastiska processer*. Teorin för stokastiska processer är ganska komplex och här ska vi bara exemplifiera kopplingen till tidsdiskreta system.

Ett exempel på en stokastisk process är följande differensekvation

$$y(k) = ay(k-1) + e(k) \tag{26}$$

där $e(k)$ är en sekvens av okorrelerade slumpmässiga variabler med medelvärde noll och varians λ. En sådan sekvens kallas *vitt brus*.

Matematiskt kan egenskaperna för vitt brus uttryckas: $E\{e(k)\} = 0$, $E\{e^2(k)\} = \lambda$, och $E\{e(k)e(j)\} = 0$ för $k \neq j$. Den sista egenskapen betyder att brusets värde vid en viss tidpunkt är helt okorrelerat med brusets värde vid andra tidpunkter!

Med olika värden på a kan olika slumpmässiga förlopp genereras, t ex om a är nära ett blir kommer två närliggande värden av y att ha en stark positiv korrelation emedan om a är nära -1 erhålls en negativ korrelation (om $y(k)$ är positiv är det en stor sannolikhet att $y(k+1)$ är negativ). Den stokastiska process som beskrivs av (26) kallas en autoreggresive process av ordning 1 eller kort AR(1) process. Det är värt att notera att det är mycket sällan som värdet på a kan bestämmas med fysikalisk modellering utan man är normalt hänvisad till empirisk modellering.

En generalisering ges av AR(N) processen som kan skrivas

$$y(k) + \sum_{p=1}^{p=N} a_p y(k-p) = e(k) \tag{27}$$