MQMO - FS20

Pascal Brunner - brunnpa7

Inhaltsverzeichnis

Kapitel 1

Einführung in das Operations Management

1.1 Operations Management in der Unternehmung

Definition:
Operations Management beinhaltet die **Gestaltung** ('Design'), den **Betrieb** ('Operation' d.h. Planen und Steuern) und die **Optimierung** ('Improvement') aller **Systeme und Prozesse**, welche beteiligt sind an der **Kreation und Bereitstellung** der **Produkte und Dienstleistungen** eines Unternehmens

- OM beschäfitgt sich damit, wie ein Unternehmen **Produkte und Dienstleistungen** *kreieren bzw. bereitstellen*
- Operations Management befasst sich mit dem Prozess von einem Delivery, Input, Process und Output. Damit wird aus dem einem Demand einen Goods oder Services erstellt. Entsprechend ist es ein zwei gleisiges Verfahren.
- OM beschäfigt sich damit, Aktivitäten des Kerngeschäfts möglichst effizient darzustellen.
- Operations bezeichnet das Kerngeschäft des Unternehmens → Alle Systeme und Prozesse, welche die Produkte und DL des UNternehmens kreieren und bereitstellen, d.h. direkt und in wesentlichem Mass an der Wertschöpfung beteiligt sind.
-

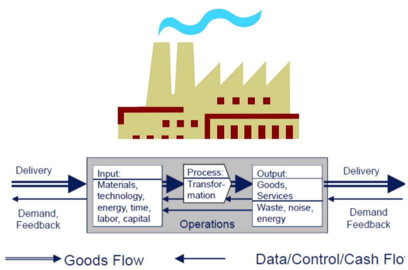


Abbildung 1.1: OM Model

1.1.1 Eingliederung von Operations

Jedes Unternehmen umfasst drei (betriebliche) Hauptfunktion:

- Marketing und Sales → Werbung, Akquise, Verkauf, Beratung
- Product/Service Development → Entwicklung, Anpassung von Produkte
- Operations → Herstellung der Produkte und Dienstleistungen, Bereitstellung für den Kunden

Zusatzfunktion:

- Engineering / Technology
- Account und Finance
- HR
- IT

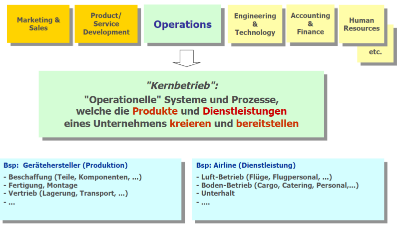


Abbildung 1.2: Operations zentrale betr. Funktion

Operations Management im Organigramm

→ Je nach Unternehmen, kann das Operations unterschiedlich verstreut sein, bspw.:

- Fertigungsbetrieb (Produktion) → typischerweise in einem Departement zusammengefasst
- Airline (Dienstleistung) → Bestimmte Bereiche der Operations-Funktion in andere Departemente ausgelagert
- Autohersteller (Kundendienst) → Operations Funktion über vers. Departemente verteilt

1.1.2 Aufgaben des Operations Management

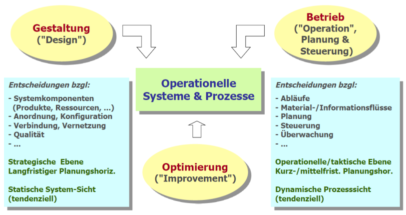


Abbildung 1.3: Die drei Aufgaben des OM

1.1.3 Wirkungsfeld des Operations Management

Durch die Industrie4.0 hat es eine zunehmende Ausdehnung des Wirkungsfelds

- Gesamtes Leistungsspektrum: Sowohl materielle **Produkte** als auch **Dienstleistungen** → Grenze zwischen Produktions- und Dienstleistungsbetrieben zunehmend fliessend
- Gesamter **Lebenszyklus** des Produkts → Beschaffung - Produktion - Vertrieb *Unterhalt* - *Entsorgung*
- Gesamte **Supply Chain** → Gesamtes Netzwerk von Unternehmen, welche an der Leistungserbringung beteiligt sind

1.1.4 Quantitatives Operations Management

Definition:
Analyse und **Optimierung** von Aufgaben des OM mit Hilfe **mathematischer Modelle und Algorithmen**
Typisches Vorgehen:

- Problem-Identifikation → kritische Teilsysteme/-prozesse identifizieren: stark ausgelastet, leistungsbestimmend, kostenintensiv, zeitkritisch
- Analyse → Sorgfältige Untersuchung: Das richtige Problem lösen
- Modellierung → Präzise Formulierung des Problems mittels mathematischer Modelle
- Optimierung → Berechnung guter Lösungen mit mathematischen Algorithmen
- Umsetzung → Übertragung / Anpassung der Resultate an das reale Problem

Decision Support: mathematische Modelle als Entscheidungshilfen (nicht 'diktatorisch')

1.1.5 Relevanz von Operations Management

- zentrale Bedeutung für Erfolg und Misserfolg eines Unternehmens
- 43 % der befragten CEO's betrachtet Operations Management als **wichtigsten Know-how Bereich** für ihre Angestellten
- OM bildet grosser Umsatzbereich weltweit im Consulting

1.2 Operations Management und Transformationsprozesse

Operations Management beschäfigt sich im Wesentlichen mit **Transformationsprozesse**

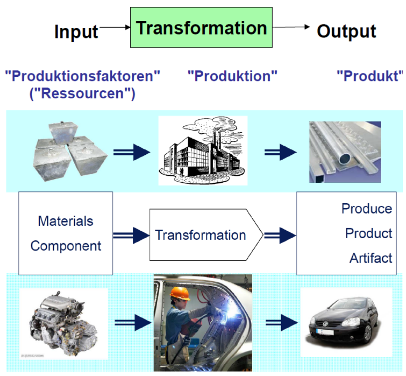


Abbildung 1.4: Schema des Transformationsprozess

1.2.1 Input

Als Input wird der **Produktionsfaktor** bzw. die Ressourcen beschrieben

- Menschliche Arbeitsleistung → objektbezogener oder dispositiver Art
- Werkstoffe → Rohstoffe, Vorprodukte, Hilfsstoffe
- Betriebsmittel - zur aktiven Nutzung → Maschinen, Anlagen, Geräte, ...
- Betriebsmittel - zur passiven Nutzung → Gebäude, ...
- Betriebsmittel - zum Verbrauch → z.B. Energie, Kühl- und Schmiermittel, ...

- Information
- Zusatzfaktoren → z.B. Versicherungen, staatliche Leistungen, ...
- Kunden (bei DL)

Dazu kommen weitere Klassifikationen:

- transformierende Ressourcen → Menschliche Arbeit, Betriebsmittel
- transformierte Ressourcen → Werkstoffe, Informationen, Kunden (bei DL)

1.2.2 Transformation

Die Transformation ist die eigentliche **Produktion**

- Physikalisch → Fertigung, Herstellung, Verbindung, ...
- Räumlich → Transporte, ...
- Zeitlich → Lagerung, ...
- Tausch → Handel, Detailhandel, ...
- Informatorisch → z.B. Telekommunikation, ...
- Physiologisch → z.B. Pflegedienste, ...

1.2.3 Output

Als Output resultiert dann das **Produkt**

- Materielle Produkte → Endprodukte, Zwischenprodukte, Abfallprodukte, ...
- Immaterielle Produkte → Dienstleistungen, Arbeitsleistungen, Information, ...
- Gemischte Produkte → unterschiedlichen Anteil an (Im)materielle Elemente

1.3 Ziele des Operations Management

Das Hauptziel von OM ist dass die Unternehmung eine möglichst grosse 'Leistung' → Performance erbringt.
Eine solche Leistung zu beschreiben und zu messen ist eine komplexe Aufgabe → **zahlreiche Leistungsmasse**.
Diese Leistungsmasse werden als KPI's (Key Performance Indicators) genannt.

1.3.1 Produktivität

Ein allgemeines, grobes Leistungsmass ist die **Produktivität** $= \frac{Output}{Input}$ Dabei ist vor allem relevant welcher *Input* bzw. *Output* in welchem *Mass* (Geldeinheit, Zeiteinheit etc.) gemessen wird

verschiedene Produktivitätsmasse

Einzelfaktoren-Produktivität Input von einzelner Ressourcen gemessen

Multifaktoren-Produktivität Input von mehreren Ressourcen gemessen (als kompatibles Mass für Input meistens Einsatzskosten in Geldeinheiten) bspw.
$Multifaktoren - Produktivitt = \frac{AusbringungsmengeeinesProdukts}{KostenArbeit + KostenMaterial + KostenEnergie}$

Output:

- Gefertigte Endprodukte	10'000 CHF	200 Stück
- Ware in Arbeit	2500 CHF	50 Stück
- Dividenden	1000 CHF	
- **Total**	**13'500 CHF**	

Input:

- Menschliche Arbeitsleistung	3000 CHF	40 Std.
- Werkstoffe	153 CHF	
- Kapital (genutzte Betriebsmittel)	10'000 CHF	
- Energie (verbrauchte Betriebsmittel)	540 CHF	
- Andere Ausgaben	1500 CHF	
- **Total**	**15'193 CHF**	

Bsp. von Einzelfaktoren-Produktivitäten:

$$\frac{\text{Total Output (CHF)}}{\text{Energie (CHF)}} = \frac{13500}{540} = 25, \qquad \frac{\text{Gefertigte Endprodukte (Stück)}}{\text{Arbeitsleistung (Std.)}} = \frac{200}{40} = 5$$

Bsp. von Multifaktoren-Produktivitäten

$$\frac{\text{Total Output (CHF)}}{\text{Arbeitsleistung (CHF) + Werkstoffe (CHF)}} = \frac{13500}{3153} = 4.28, \qquad \frac{\text{Gefertigte Endprodukte (CHF)}}{\text{Arbeitsleistung (CHF) + Werkstoffe (CHF)}} = \frac{10000}{3153} = 3.17$$

Abbildung 1.5: Numerisches Beispiel für die Produktivität

Nutzen und Grenzen von Produktivitätsmasse

- relatives Mass (nicht absolut!) → aussagekräftig im Vergleich (andere Abteilungen, über Zeit, Konkurrenten)
- Produktivitätsmasse als operative Ziele meist ungeeignet

1.4 Entscheide des Operations Management

Entscheidungsbereiche (Aufgaben)

- Gestaltung ('Design')
- Betrieb ('Operation') → Planung (Planning) und Steuerung (Execution Control)

Abbildung 1.6: Entscheidungsbereiche im OM

Entscheidungsebene (entsprechend Tragweite)

- strategisch → langfristig
- taktisch → mittelfristig
- operativ → kurzfristig

TABLE 1.1 Differentiating Factors of the Three Decision Categories

FACTOR	STRATEGIC PLANNING	MANAGEMENT CONTROL (TACTICAL PLANNING)	OPERATIONAL CONTROL
Purpose	Management of change, resource acquisition	Resource utilization	Execution, evaluation, and control
Implementation instruments	Policies, objectives, capital investments	Budgets	Procedures, reports
Planning horizon	Long	Medium	Short
Scope	Broad, corporate level	Medium, plant level	Narrow, job shop level
Level of management involvement	Top	Middle	Low
Frequency of replanning	Low	Medium	High
Source of information	Largely external	External and internal	Largely internal
Level of aggregation of information	Highly aggregated	Moderately aggregated	Detailed
Required accuracy	Low	Medium	High
Degree of uncertainty	High	Medium	Low
Degree of risk	High	Medium	Low

Abbildung 1.7: Entscheidungsebene im OM

Entscheidungselemente

- Produkte
- Potentiale (Kapazitäten)
- Prozesse

Tragweite der Entscheidung / Elemente	Strategisches Produktionsmanagement	Taktisches Produktionsmanagement	Operatives Produktionsmanagement
Produkt- und Programmgestaltung	Festlegung der Produktfelder	- Konkretisierung der Produktfelder nach Art und Qualität - Festlegung der Breite und Tiefe des Produktionsprogramms - Planung neuer Produkte - Verbesserung vorhandener Produkte	Festlegung des Produktionsprogramms hinsichtlich Art und Menge
Potentialgestaltung	- Kapazitätsdimensionierung - Festlegung der Kapazitätsarten - Langfristige Rohstoffversorgung	- Personal- und Maschinenausstattung Technologieeinsatz - Festlegung der Bestellpolitik - Bestimmung optimaler Losgrößen	Bereitstellung der Produktionsfaktoren für ein gegebenes Produktionsprogramm - Beschaffung von Repetierfaktoren - Reservierung vorhandener Anlagen - Einsatz der Mitarbeiter
Prozeßgestaltung und -steuerung	Festlegung des generellen Prozeßablaufes in der Produktion (Produktionstyp)	- Festlegung des technologischen Verfahrens - Festlegung der innerbetrieblichen Standorte	Sicherstellung des optimalen Ablaufes des Produktionsprozesses bei gegebenem Produktionsprogramm (Maschinenbelegung, Reihenfolgeplanung)

Abbildung 1.8: Entscheidungselemente im OM

1.4.1 Unternehmenstrategie

Wird grundsätzlich in die vier Bereiche unterteilt:

- Quality
- Costs
- Delivery
- Flexibility

Diese vier Bereiche stehen natürlich in einem gewissen Masse in einem Zielkonflikt. Da sehr hohe Qualität sehr selten mit sehr niedrigen Kosten hereingehen usw.

1.5 Arbeitsdefinitionen

System

- Menge von Objekten (Entitäten, Systemelementen), welche zu einem bestimmten Zweck miteinander verbunden sind und gegenseitig interagieren
- Beschreibung eines Systems oft anhand eines (evtl. zeitabhängigen) Systemzustands
- *Systemanalyse* Eher statische (zeitunabhängige) Betrachungsweise: 'Was sind die Systemelemente, wie sind sie miteinander verbunden und wie interagieren sie?

Prozess

- Ablauf einer Folge von zusammenhängenden Aktionen (Aktivitäten, Operationen, Handlungen,...)
- Folge von Zustandsänderungen eines Systems
- *Prozessanalyse:* Eher dynamisch (zeitabhängige) betrachtet: 'Was sind die zeitlichen Abläufe / Handlungsfolgen bzw. die zeiitlichen Veränderungen des Systemzustands?'

Planung

- Plan: Handlungsabsicht, gedankliche Vorwegnahme zukünftiger Handlungen
- Planung: Entwicklung eines Plans (zukunftsgerichtet): Festlegen von Zielen (Sollwerten) und von Methoden zur Erreichung dieser Ziele

Steuerung

- Umsetzungeines Plans (gegenwartsbezogen): Veranlassung von Handlungen, Kontrolle von ISTwerten gegenüber SOLLwerten, Anpassung des Plans ('Regelkreis')

Kapitel 2

Einführung in die Prozess-Analyse

2.1 Qualitative Prozessanalyse

Für die qualitative Prozessanalyse gibt es eine Vielzahl an Notationen (bspw. BPM), Begrifflichkeiter, jedoch **keine** klare einheitliche bzw. normierte Terminologie.
Wobei der Begriff *Business Process* ein betrieblicher Ablauf darstellt:

- Menge von Teilprozessen (Aktivitäten, Vorgängen, Aktionen, Tätigkeiten, ...), welche in bestimmten logischen und zeitlichen Zusammenhang stehen
- **EN Definition:** *A business process is a colleciton of interrelated work tasks, initiated in response to an event, that achieves a specific result for the customer of the process*

Ziele:

- Analyse bestehender Prozesse ('As-Is') → identifizieren, Verstehen, Dokumentieren Verbesserung bestehender Prozesse, Gestaltung neuer Prozesse ('To-Be') → Aufdecken von Optimierungspotential, Umsetzung in neue Abläufe

Methodik:

i Sprachliche Beschreibung der Prozesse

ii Graphische Darstellung in Form von Prozessdiagrammen

2.1.1 Grundlagen der Prozessdiagramme

Prozessdiagramme sind eine grafische Darstellung eines Prozesses, welche Teilprozesse und deren Zusammenhänge abbilden (bspw. mit Symbole und Pfeile)
Für diese grafische Darstellung gibt es viele unterschiedliche notationen (bspw. Anzahl Symbole, Pfeiltypen, Ziel, Spezifikation). Es ist elementar diese Notation zu verstehen, damit man das Diagramm korrekt liest, ansonsten können missinterpretationen folgen. Beispiele von unterschiedlichen Arten:

- Hierarchische Strukturierung von Prozessen → Baukasten-System, Detailgrad abhängig von Anwendung und Zielsetzung
- Questions and Facts (bspw. IT-Anwendung)
- nesting structure

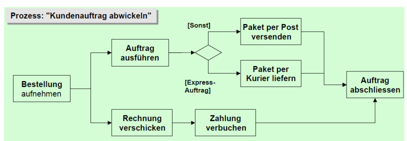


Abbildung 2.1: Beispiel eines Prozessdiagrammes für die Kundenauftrag Abwicklung

Prozessdiagramm-Typen

→ Es gibt zahlreiche Diagrammsprachen (Modellierungssprachen).
Definition:

- **Syntax:** *Konstruktionsregeln für zulässige Diagramme (Symbole, Verbindungen, ...)*
- **Semantik:** *Erklärung der Bedeutung (Interpretation) der verschiedenen Sprachelemente*
- → In der Praxis häufig nicht formal exakt definiert

elementare Grundtypen: → Unterscheidung bezüglich Interpretation der Pfeile
Aktivitätsdiagramme (Activity Diagrams, Workflow Diagrams, ...):
→ Pfeil bedeutet Zeitliche Präzedenz (genauer 'Sequenzfluss')

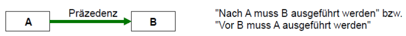


Abbildung 2.2: Beispiel eines Aktivitätsdiagramme

Flussdiagramme (Flow Diagram, Material Flow Diagram, ...):
→ Pfeil bedeutet Fluss von OBjekten von A nach B (Material, Aufträge, ...)

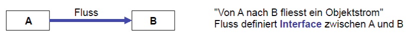


Abbildung 2.3: Beispiel eines Flussdiagramms

Entwicklung von Prozessdiagramme

- **Wichtig:** Klare Zielsetzung/Fragestellung bei der Prozessanalyse
- Schwierigkeiten → Identifikation/Abgrenzung der Prozesss und Detaillierung/Verknüpfung der Subprozesse
- Erstellen von guten Diagrammen ist eine 'Kunst'
- Gefahr: Paralysis By Analysis → Zu hohe, unangemessene Ansprüche an Detaillierungsgrad ⇒ Modellierung endlos

Hilfreiche Anwendungtipps:

- Detail nur soweit wie Zielsetzung
- Diagramme so einfach wie möglich halten
- Diagramme sollten möglichst selbsterklärend sein

2.1.2 Aktivitätsdiagramme

Syntax

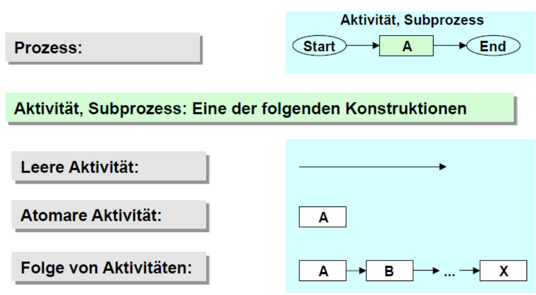


Abbildung 2.4: Syntax der Aktivitätsdiagramme

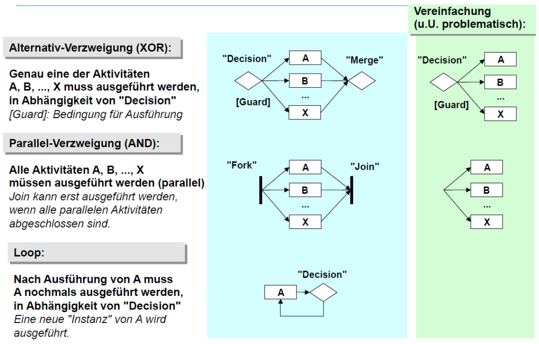


Abbildung 2.5: Syntax der Aktivitätsdiagramme II

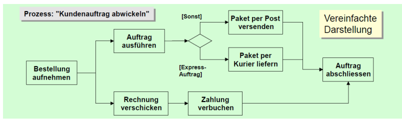


Abbildung 2.6: Vereinfachte Darstellung eines Aktivitätsdigramm für den Kundenauftrag abwickeln

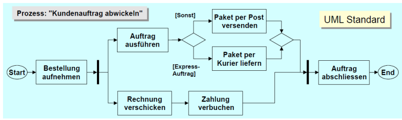


Abbildung 2.7: UML Darstellung eines Aktivitätsdigramm für den Kundenauftrag abwickeln

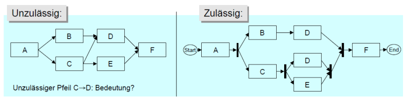


Abbildung 2.8: un- bzw zulässig eines Aktivitätsdigramm

Semantik

Jeder *Weg* (Routing) vom Start- zum Endkonoten beschreibt möglichen *Verlauf* des Prozesses

2.1.3 Swimlane-Diagramme

Gilt als Erweiterung der Aktivitäts-Diagramme - wobei eine zusätzliche Information (der Akteur / die Rolle) ersichtlich ist → *Wer führt eine Aktivität durch (oder: ist zuständig dafür...)*
ohne Swimlane: *Was wird in welcher Reihenfolge ausgeführt?*
mit Swimlane: *Was wird von wem in welcher Reihenfolge ausgeführt?*
typische Akteure: Person, Personengruppe, Jobfunktion, Organisationseinheit, Maschine, Einrichtung, Computer

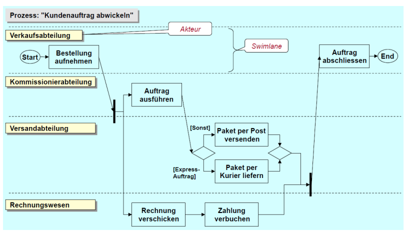


Abbildung 2.9: Beispiel eines Swimlane-Diagramm

2.1.4 Flussdiagramme

wichtig! Keine normierte Terminologie

Semantik

- Pfeile stellen Flüsse von Objekten dar (Material, Aufträge, ...)
- Flüsse definieren Schnittstellen zwischen Subprozessen (Aktivitäten): Input, Output
- Flussobjekte werden im Subprozess transformiert (Physikalisch, örtlich, zeitlich)

Syntax

Für jeden Flusstyp, d.h. jede unterschiedlcihe Art von Flussobjekten: Diagramm bildet ein sogenanntes **S-T-Netzwerk**:

- zusammenhängendener Graph (Knoten: Aktivitäten, Pfeile: Flüsse)
- Genau ein Startknoten: **Quelle (S: Source)** → generiert FLussobjekte, nur austretende Pfeile
- Genau ein Endknoten: **Senke (T: Target)** → konsumiert Flussobjekte, nur eintretende Pfeile
- Alle anderen Zwischenknoten brauchen mind. 1 ein- und 1 austretender Pfeil
- Manchmal werden Quelle und Senke auch weggelassen

Unterschied zu Aktivitätsdiagrammen

- Keine Parallelverzweigung
- Jede Verzweigung ist eine Alternativ-Verzweigung → Bei Verzweigung häufig Prozentangabe für Flussanteil
- **Routing** Möglicher Weg eines Flussobjekts von der Quelle zur Senke

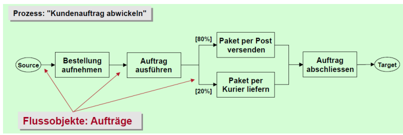


Abbildung 2.10: Beispiel eines Flussdiagramm

Zu beachten

- Immer Typ (Auftrag, Komponente, Gewichtseinheiten, ...) und Masseinheit (Anzahl, Kg, Stunden, ...) der Flussobjekte spezifzieren
- Flüsse verschiedener Objekttypen/-einheiten können nicht zusammengeführt werden
- vers. Flusstypen im gleichen Diagramm graphisch unterscheiden (bspw. vers. Pfeildarstellung)

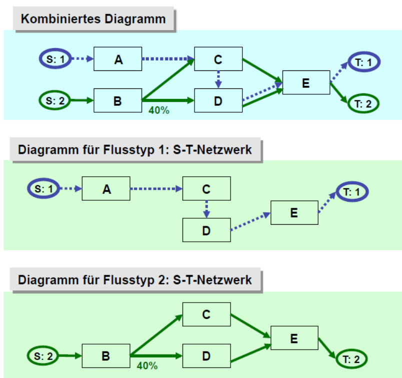


Abbildung 2.11: Beispiel eines Flussdiagramm mit zwei Flusstypen

Puffer

- Puffer sind spezielle Subprozesse (Aktivitäten), welche zur 'zeitlichen Transformation' (Lagerung) dienen (→ Zwischenspeicher, Lager, Warteraum etc)
- da spezielle Bedeutung häufig eigenes Symbole
- Mehrstufige Prozesse sind häufig gepuffert
- Dienen dazu **Blocking** (Blockierung) und **Starving** (Aushungern) zu vermeiden

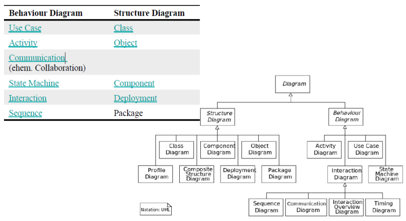
2.1.5 UML und Standards

einige Standards

- Integrated DEFinition → IDEF ⇒ Sammlung verschiedener Modellierungssprachen
- Business Process Modeling Notation → BPMN
- Unified Modeling Language → UML ⇒ Standard in Software-Entwicklung; Sammlung von zahlreichen versch. Diagrammtypen inkl. Spezifikation

UML

In UML sind folgende Diagramme am wichtigsten:



Behaviour Diagram	Structure Diagram
Use Case	Class
Activity	Object
Communication (ehem. Collaboration)	
State Machine	Component
Interaction	Deployment
Sequence	Package

Abbildung 2.12: Übersicht der wichtigsten UML Diagramme

Object Diagram
Class Diagram

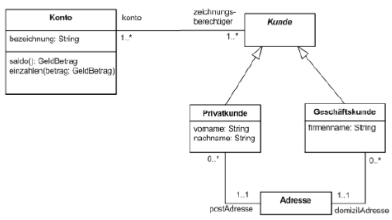


Abbildung 2.13: Class Diagram nach UML

Sequence Diagram

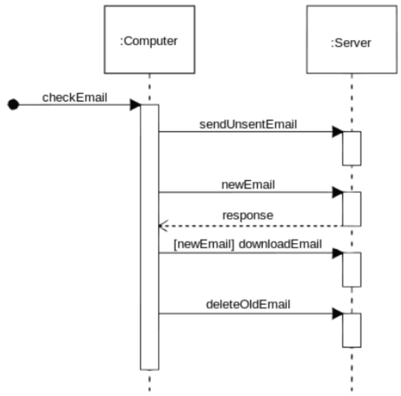


Abbildung 2.14: Sequence Diagram nach UML

Activity Diagram

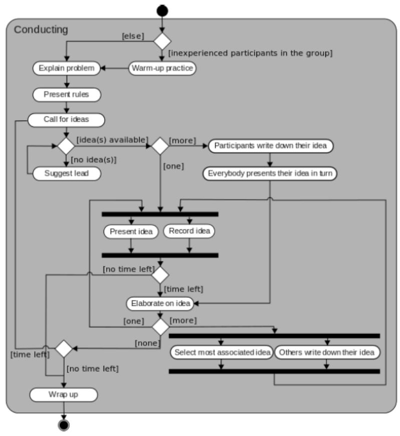


Abbildung 2.15: Activity Diagram nach UML

UseCase Diagram

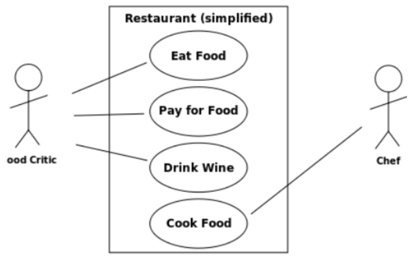


Abbildung 2.16: UseCase Diagram nach UML

State Machine Diagram

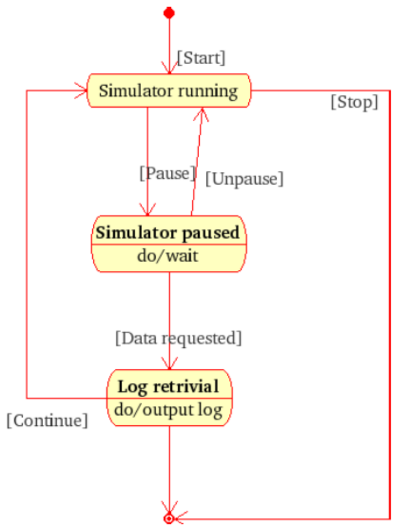


Abbildung 2.17: State Machine Diagram nach UML

Communication Diagram

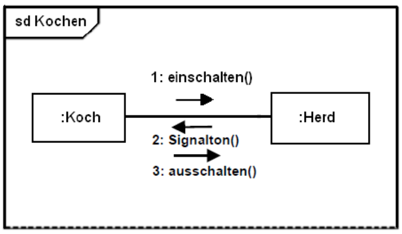


Abbildung 2.18: Communication Diagram nach UML

Interaction Diagram

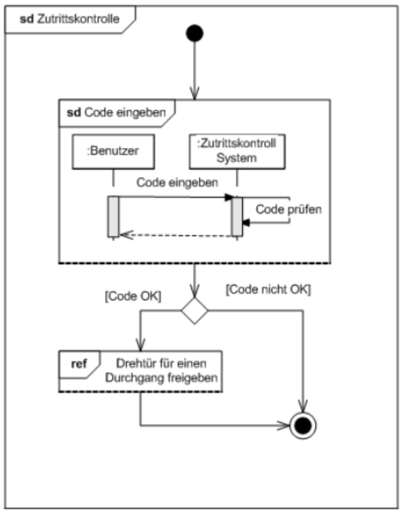


Abbildung 2.19: Interaction Diagram nach UML

Deployment Diagram

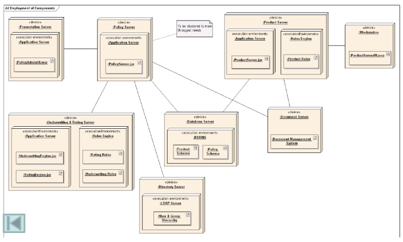


Abbildung 2.20: Deployment Diagram nach UML

Component Diagram

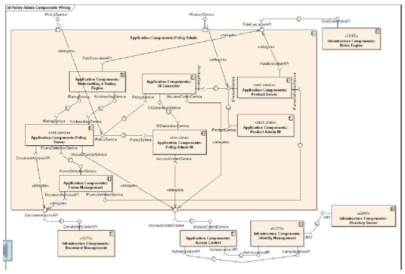


Abbildung 2.21: Component Diagram nach UML

Package Diagram

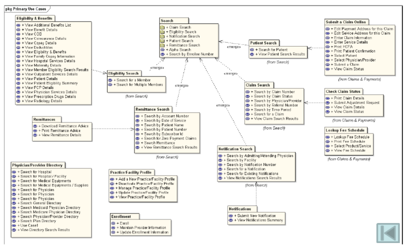


Abbildung 2.22: Package Diagram nach UML

2.2 Quantitative Prozessanalyse

2.2.1 Formel von Little

Betrachte einzelnen Prozess (Subprozess) als Blackbox, welche Input zu Output transformiert ($\rightarrow$ Transformationsprozess).
Annahme: Nur ein Typ von Flussobjekten (Flusseinheit) fliesst durch den Prozess.

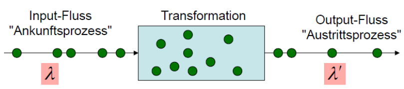


Abbildung 2.23: Abbildung des Flusses nach Little

Gleichgewichtsbedingung

Ein Prozess ('System') ist im **Gleichgewicht** ('stabil') wenn $\lambda = \lambda'$ gilt.
Wobei $\frac{1}{\lambda}$ die mittlere Zwischenankunftszeit beschreibt

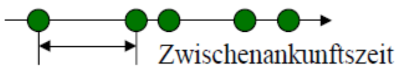


Abbildung 2.24: Abbildung der Zwischenankunftszeit

Die 3 zentralen Leistungskennzahlen eines (stabilen) Prozesses

λ, λ': mittlere Ankunfts-/Austrittsrate ('mean arrival / exit rate') $\Rightarrow \frac{Anzahleintretende/austretendeFlusseinheit}{Zeiteinheit}$ in der Masseinheit [FE/ZE]
N: mittlere Anzahl Objekte im System ('mean number of objects in system') in der Masseinheit [FE]
W: mittlere Aufenthaltszeit eines Objekts im System ('mean sojourn time') in der Masseinheit [ZE]

Wobei *mittlere* über langen Zeitraum bedeutet.

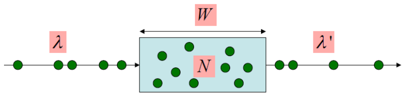


Abbildung 2.25: Abbildung eines stabilen Prozesses

Die Formeln

In der quantitative Prozessanalyse ist die Formel von Little ein fundamentaler Zusammenhang, welcher unter sehr allgemeinen Bedingungen gültig ist. Es wird auch das Ohmsche Gesetz in der Prozessanalyse genannt.

$$N = \lambda * W \tag{2.1}$$

Doch auch im Operations Management hat es eine wichtige Bedeutung:

$$WIP = TH * CT \tag{2.2}$$

WIP: (mean) Work in Process → mittlere Ware-In-Arbeit ('WIA')
TH: (mean) Throughput → mittlerer Durchsatz, mittlere Produktionsrate
CT: (mean) Cycle Time → mittlere Durchlaufzeit ('DLZ')

Anwendung in Operations Management

- Anwendbar auf bel. Prozesse (Bereiche) in einem Betrieb → Ganzer Betrieb, Teilbetrieb, Arbeitsstation, Arbeitsplatz, Bedienschalter, Maschine, ...
- Allgemeine relation zwischen den 3 Kenngrössen (WIP, TH, CT) → Falls zwei Grössen bekannt sind, kann dritte Grösse berechnet werden
- Reduktion von Durchlaufzeiten (bei konstanter Produktionsrate) → DLZ wächst/sinkt proportional zu WIA $CT = \frac{1}{TH} * WIP$
- Ermittlung Durchlaufzeit in Praxis → DLZ ist sehr aufwendig zu messen, jedoch lässt sich dies durch die beiden anderen Kenngrössen WIP und Produktionsrate ermitterln $CT = \frac{WIP}{TH}$
- Messung der WIP in Form von Zeitaufwand → Little's Formel kann als Umrechnung der Masseinheit WIP gesehen werden

2.2.2 Kapazität und Auslastung, Bottleneck

Begriff der Kapazität

- Kapazität als mengemässiges Leistungsvermögen, welches immer bezüglich eines **enzigen Objekttyps (Produkts)** gemessen wird.

- **Definition:** *Kapazität einer Ressource (eines Prozesses) bezüglich eines bestimmten Objekttypes (Produkts): Die maximal mögliche Ausbringsungsmenge (Output) fpr diesen Objekttyp (Produkt) in einem bestimmten Zeitraum*

- Kapazität μ = maximal möglicher (mittlerer) Throughput
- Masseinheit: [Anzahl Objekte / Zeiteinheit]
- per Definition gilt $\lambda \leq \mu$ bzw. TH $\leq \mu \Rightarrow \lambda =$ TH die Ankrunftstrate bzw. den Throughput bezeichnet
- minimale (mittlere) Aufenthaltsdauer ('mittlere Bearbeitungszeit') $= \frac{1}{\mu}$
- Falls mehrere Objekttypen $i \in I$ durch einen Prozess fliessen $\rightarrow$ Kapazität μ_i = maximal möglicher mittlerer Throughput bzgl. i

Begriff der Auslastung

- Auch Auslastungsgrad, Kapazitätsauslastungsgrad genannt
- Auslastung einer Ressource (Prozesses): Auslastung $= \frac{GenutzeKapazitt}{NutzbareKapazitt}$
- Masseinheit: einheitslos (bspw. %)
- **Bottleneck** eines Routings:Teilprozess mit grösster Auslastung
- Auslastung p $= \frac{\lambda}{\mu} = \frac{Throughput}{NutzbareKapazitt}$ wobei immer $p > 1$ gilt
- bei mehreren Objekttypen $i \in I$: Auslastung bzgl. Objekttyp (Produkt) i $p_i = \frac{\lambda_i}{\mu_i}$
- bei mehreren Objekttypen $i \in I$: Gesamtauslastung $p = \sum_{i \in I} p_i$

Kapitel 3

Elemente der Warteschlangentheorie

3.1 Einführung und Grundlage

In Operations Management geht es darum die Ressourcen so dimensionieren, dass

- Warteschlange klein, bzw. selten sind, ein Kunde im MIttel wenig wartet, ein gewisser *Servicegrad* gesichert ist
- Ein *Kostenoptimum* erzielt wird

→ wird auch als Wartephänomen zusammengefasst, welches in OM allgegenwertig ist

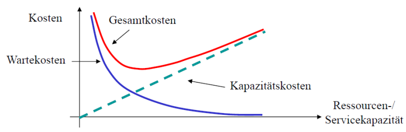


Abbildung 3.1: Abbildung eines Kostenoptimum

- Warteschlangenphänomene entstehen mit dem Verbrauch von beschränkten Ressourcen
- Ressource (physisch, logisch, menschlich) wird benötigt um Tätigkeit an Entitäten auszuführen, welche allgemein mit *Kunden* bezeichnet werden
- Auch wenn im Mittel genügend Ressourcen vorhanden, können Konflikte entstehen → Wartephänomene für Kunden
- Warteschlangenphänomene können quantiativ untersucht werden (einfachen Fälle: Queueing Theorie; komplexe Fälle: Simulationsmodelle) ⇒ Bei beidem handelt es sich um stoachastische Modelle, d.h. gewisse Daten sind vom Zufall bestimmt und man will sie als Zufallsvariablen Modellierungssprachen
- Beide sind Modelle, die zeitliche Prozesse beschreiben
- Queueing Theorie ist das analytische Werkezug für die Modellierung

3.1.1 Grundbegriffe der Warteschlangentheorie

Eine Wartschlange ('Queue') ist ein stoachastisches (= zufälliges) System, das sich aus folgenden schematischen Elementen zusammensetzt.

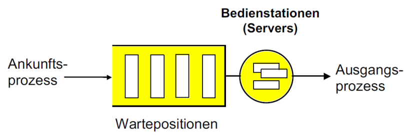


Abbildung 3.2: Abbildung der Warteschlangentheorie

Ankunftsprozess

- Die Ankünfte der Kunden bilden einen stoachastischen Prozess → **Zwischenankunftszeit ('Interarrival Time', IAT)**
- Annahme: IAT ist unabhängig und gleichverteilt
- Notation **M**: Exponential-Verteilung (Markovian)
- Notation **D**: konstante Verteilung (Deterministic)
- Notation **G**: Beliebige Verteilung (General)

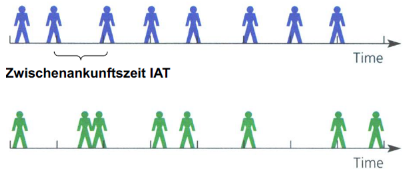


Abbildung 3.3: Abbildung der Zwischenankunftszeit

Absolute Variabilität: Standardabweichung
Relative Variabilität: Variationskoeffizient (CV): $CV = \frac{Standardabweichung}{Erwartungswert}$

Warteschlangen-Disziplin

Die Disziplin legt fest in welcher Reihenfolge der Zugriff stattfindet auf die Ressourcen
FIFO bzw. FCFS (First in, first out bzw. first come, first served) → Kunden werden in der Reihenfolge der Ankunft bedient. ⇒ Allgemeines Default
LIFO bzw. LCFS (Last in, frist out bzw. last come, first served) → Kunden werdden in umgekehrter Reihenfolge der Ankunft bedient
SIRO (Service in random order) → Kunde wird zufällig aus der Wartschlange ausgewählt
PS (Processor Sharing) → Kunden teilen sich eine Bedienungsstelle. bei **n Kunden** erhält jeder $\frac{1}{n}$ der vorhandenen Ressourcen
RR (Round Robin) → Bedienung der Kunden im Kreis herum. Jeder erhält nacheinander ein Quantum σ. Falls noch mehr benötigt wird, stellt er sich wieder in die Warteschlange

Bei der **Preemption** werden die Kunden im Vorfeld nach Priorität in unterschiedlichen Klassen aufgeteilt.

Bedienung der Kunden

Annahme: Die **Bedienzeit** (Service Time, ST) für die Bedienung eines Kunden sind unabhängige und gleichverteilte Zufallsvariablen.
Dabei entspricht die Anzahl Kunden, welche *gleichzeitig* bedient werden können der **Anzahl Bedienstationen** der Warteschlange (→ sämtliche Bedienstationen sind identisch)

Noation von Kendall

Wartschlange wird mit einer symbolischen Formel beschrieben: **A/B/s**
Wobei:
A → Verteilung der Zwischenankunftszeiten ($A \in [M, D, G]$)
B → Verteilung der Bedienzeiten ($B \in [M, D, G]$)
s → Anzahl Bedienstationen ($s \in [1, 2, .., n]$)

Folgende Annahmen gelten:

- Die Menge (Population) der Klienten, welche in das System eintreten können, ist unbeschränkt

- die Anzahl der Klienten, welche sich inder Wartschlange befinden können, ist unbeschränkt
- die Wartschlangen-Disziplin ist FIO

Erweiterte Form der Notation von Kendall → A / B / s / K / P / DS
K → Kapazität der Warteschlange (Anzahl Warteplätze + Anzahl Bedienstationen)
P → Grösse der Population
DS → Warteschlangen-Disziplin

⇒ Fall einer beschränkten Kundenpopulation → Ankunftsprozess hängt von Anzahl Kunden im System abbilden
⇒ Fall einer beschränkten Aufnahmekapazität → Was passiert mit Ankommende, wenn Warteschlange voll ist? Kommen sie wieder oder nicht?

Leistungsmasse

N → Anzahl Kunden im System (an Warteposition oder in Bedienung)
N_q → Anzahl Kunden in Warteschlange
W → Aufenthaltszeit eines Kunden im System
W_q → Wartezeit eines Kunden in Warteschlange
⇒ Diese Grössen sind Zufallsvariablen, deren Erwartungswert, Verteilung, Quantile (Schwellenwert) etc. abgeschätzt werden und sind im Allgemeinen von der Zeit t abhängig. Für die Abschätzung ist es einfacher wenn das System sich in einem Gleichgewichtszustand befindet.

Beschreibende Parameter

λ	(mittlere) Ankunftsrate
$1/\lambda$	(mittlere) Zwischenankunftszeit
μ	(mittlere) Bedienrate (entspricht "Kapazität" einer Arbeitsstation)
$1/\mu$	(mittlere) Bedienzeit
s	Anzahl Arbeitsstationen

Davon abgeleitete Grösse:

ρ	(mittlere) Auslastung (auch: Auslastungsgrad, Ausnutzungsgrad, Verkehrsintensität, ...)

Es gilt:

$\rho = \frac{\lambda}{s\mu}$	allgemein bei $s \geq 1$ Arbeitsstationen
$\rho = \frac{\lambda}{\mu}$	speziell bei $s = 1$ Arbeitsstation

Abbildung 3.4: Abbildung der beschreibenden Parameter

Leistungskennzahlen

$E[N]$	(mittlere) Anzahl Kunden im System
$E[N_q]$	(mittlere) Anzahl Kunden in Warteschlange
$E[W]$	(mittlere) Aufenthaltszeit eines Kunden im System
$E[W_q]$	(mittlere) Aufenthaltszeit eines Kunden in Warteschlange ("Wartezeit")

p_i	Wahrscheinlichkeit, dass genau $i \geq 0$ Kunden im System sind d.h. $p_i = P[N = i]$

Abbildung 3.5: Abbildung der Leistungskennzahlen

Grundlegende Zusammenhänge

$$E[W] = \frac{E[N]}{\lambda} \qquad \text{Formel von Little}$$

$$E[W_q] = E[W] - \frac{1}{\mu}$$

$$E[N_q] = \lambda \cdot E[W_q] \qquad \text{Formel von Little}$$

Abbildung 3.6: Abbildung der grundlegende Zusammenhänge

→ Falls E[N] gegeben ist, können die anderen drei Leistungsmasse daraus berechnet werden

Für das Operations Management (in Fertigung)

- System: Produktionseinheit, Werkstatt, Fertigungszelle
- Bedienungsstelle: Maschinen, Arbeits-Stellen/-Stationen
- Anzahl Kunden im System: Ware in Arbeit
- Anzahl wartender Kunden: bspw. auf Verarbeitung von einer Maschine wartende Stücke
- Wartezeit eines Kunden: Warte-/Liegezeit eines Stücks
- Im System verbrachte Zeit eines Kunden: Durchlaufzeit
- Auslastungsgrad der Bedienungsstelle: Kapazitätsauslastungsgrad (der Maschine, Arbeitsstelle)

3.2 Discrete Event Simulation

Betriebliche Systeme und Prozesse zeichnen sich dadurch aus, dass die Systemzustände überlicherweise diskret sind, d.h. durch diskrete Zahlen beschrieben werden können. bspw. (Anzahl offener Bestellungen, Anzahl Teile im Lager, Auftrag in Bearbeitung)
→ Zustandsgrössen ändern sich sprunghaft zu einem bestimmten Zeitpunkt

3.2.1 Modellierungsansätze

Kontinuierliche Zeitmodellierung

Charakteristika: Vorgängig festgelegte Zeitachse; Zustandsveränderungen zwischen aufeinanderfolgenden Zeitpunkten. Es werden eher Flüsse oder stetige Prozesse modelliert.

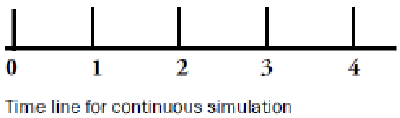


Abbildung 3.7: Abbildung einer kontinuierlichen Zeitmodellierung

Diskrete Ereignis Modellierung

Charakteristika: Zustandsänderungen geschehen auf Grund von stochastischem Proezss. Die Zeitpunkte werden durch die Ereignisse der Zustandsänderung bestimmt. Es werden eher Objekte, einzeln oder in Gruppen modelliert. DIe Objektzustände, die über gewisse Zeitintervalle bis zur nächsten Zustandsänderung unverändlich sind charakterisieren das Modell.

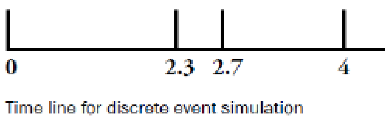


Abbildung 3.8: Abbildung einer diskreten Ereignis Modellierung

Diskrete Raten Modellierung

Charakteristika: Hybrid aus kontinuierlicher und ereignisorientierter Modellierung. Es werden Flüsse, die durch (bspw. Durchsatz-) Raten charakterisiert werden modelliert. Auf Grund von diskreten Ereignissen ändern sich die Raten.

Übersicht der Methoden

Chapter03 - Ab Slide 25

3.2.2 Ablauf einer Simulationsstudie

1. Problembeschreibung

- Definition des genauen Problems, das untersucht werden soll
- Schriftliche Aussage über das Ziel der Problemstellung

2. Modellierung des Systems und der Prozesse

- Was soll modelliert werden?
- Wie sind die Prozesse beschaffen?

3. Datensammlung

- Welche realen Daten des Systems werden wie verwendet?

4. Bau des Simulationsmodells

- Welche Prozesse sollen berücksichtigt werden?
- Wie werden diese modelliert?
- Welche Effekte werden vernachlässigt?
- Wie wird die Variablität beschrieben?

5. Verfikation

- Macht das Modell, was es machen soll?

6. Validierung

- Ist das Modell ein korrektes Abbild der Realität?
- Beschreibt das Modell die Wirklichkeit genügend gut?

7. Durchführung von Experimenten

- Verschiedene Simulationsruns, um bestimmte Fragen zu beantworten

8. Ergebnisanalyse und Interpretation

- Analyse der Ergebnisse (i.d.R. mit statistischen Auswertungen)

9. Dokumentation

10. Umsetzung

3.2.3 Algorithmus

Der Simulationsalgorithmus

- Eventliste L wird so geordnet, dass der EVent mit der kleinsten Ausführungszeit t_k an erster Stelle steht. Dieser Event ist der Event, das als nächstes bearbeitet wird
- Die Bearbeitung dieses Events zur Systemzeit t_{trigg} (t_{trigg} Ausführungszeitpunkt des Events) $\rightarrow$ ändert den Systemzustand, $\rightarrow$ löst neue Events aus, $\rightarrow$ löscht pending Events
- Updaten der Even-List L $\rightarrow$ der erste Event (welcher gerade bearbeitet wurde) wird entfernt, $\rightarrow$ Eventliste neu sortieren, so dass der oberste Event dasjenige ist, das als nächstes ausgeführt wird.

3.3 Markov Eigenschaften

Dem Modell M/M/1 liegt die **Markov-Kette**.
Es handelt sich um einen Prozess mit abzählbar vielen Zuständen und gedächtnislosen Übergangswahrscheinlichkeiten.

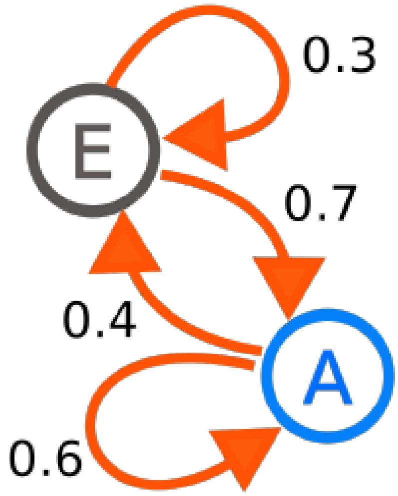


Abbildung 3.9: MM1 Modell

3.3.1 Markov-Kette

Die wichtigste Eigenschaft von Markov Ketten ist, dass ihr stochastisches Verhalten durch Übergangswahrscheinlichkeiten beschrieben werden kann.

$$P[X(t_{k+1}) = x | X(t_k) = x'] \tag{3.1}$$

für Zustände X(t_k) zum Zeitpunkt t_k mit Zustandswerten x, x' und $t_k \leq t_{k+1}$
Wenn diese Übergangswahrscheinlichkeiten, sowie eine Häufigkeitsverteilung des Anfangzustandes gegebensind, ist es möglich, die Wahrscheinlichkeit jedes Zustands an jedem Zeitpunkt zu bestimmen.

Eigenschaft stoachastischer Prozesse

→ Eingangswahrscheinlichkeit aller Zustände X(0) = x zur Zeit $t_k = 0$ müssen in Summe 1 ergeben.

$$\sum P[x(T_k = 0) = x] = 1 \tag{3.2}$$

→ Alle von einem Zustand $X(t_k) = x$ ausgehenden Übergangswahrscheinlichkeiten müssen 1 ergeben:

$$\sum P[X(t_k + 1) = x' | X(t_k) = x] = 1 \tag{3.3}$$

für Zustände X($t_k + 1$) zum Zeitpunkt t_k

Regel für die Gesamtwahrscheinlichkeit

In Markov-Ketten wollen wir die Gesamtwahrscheinlichkeit $p(x', x)$ für einen Übergang vom Zustand x in x' betrachten, unabhängig davon welches Ereignis diesen Zustandsübergang tatsächlich ausgelöst hat.
Wir wenden also die Regel für die Gesamtwahrscheinlichkeit an:

$$p(x'|x) = \sum_i p(x'|x, i) * p(i|x) \tag{3.4}$$

wobei p(i,x) die Wahrscheinlichkeit ist, dass Ereignis i im Zustand x eintritt; und p(x'|x,i) die Wahrscheinlichkeit, dass mit Ereignis i im Zustand x daraufhin Zustand x' eintritt.

Die Matrix der Übergangswahrscheinlichkeiten

Die Informationen der Übergangswahrscheinlichkeiten in einer zeitdiskreten Markov-Kette wird in Matrixschreibweise adäquat beschrieben.
Wir definieren die *transition probability matrix* **P**, welche aus den Koeffizienten p_{ij} besteht:

$$P \equiv [p_{ij}], i, j = 0, 1, 2, .., n \tag{3.5}$$

wobei n die Anzahl Zustände im Zustandsraum angibt.

Dabei gilt: Alle Elemente jeder i-ten Zeile dieser Matrix, i=1,2,... müssen sich immer zu 1 addieren.

$$\begin{bmatrix} p_{00} & p_{01} & .. & p_{0j} \\ p_{10} & p_{11} & .. & p_{1j} \\ p_{20} & p_{21} & .. & p_{2j} \end{bmatrix}$$

State transition Diagram

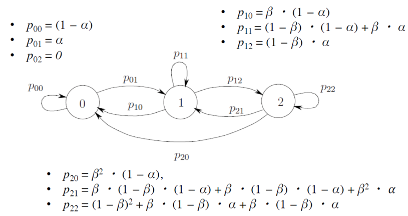


Abbildung 3.10: Das State Transition Diagram

3.4 Das Model M/M/1

Die Kunden gelangen in die Schlange gemäss einem *Poisson-Prozess*, d.h. die **Zwischenankunftszeiten** sind voneinander unabhängig, identisch **exponential-verteilte (iid)** Zufallsgrössen.
Die Dichte ist

$$f_A(t) = \lambda e^{-\lambda t} \tag{3.6}$$

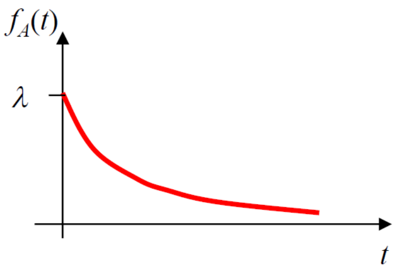


Abbildung 3.11: Funktion MM1

→ λ ist die mittlere Ankunftsrate
→ $\frac{1}{\lambda}$ ist die mittlere Zwischenankunftszeit (d.h. der Erwartungswert der durch f_A definierten Verteilung)

Eine **einzige** Bedienungsstelle kümmert sich um diese Kunden und zwar in der Reihenfolge FIFO. Die jeweiligen Bedienungszeit ist eine exponential-verteilte Zufallsgrösse mit Dichte:

$$f_s(t) = \mu e^{-\mu t} \tag{3.7}$$

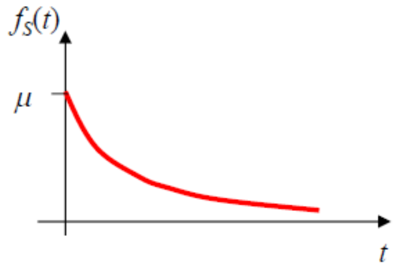


Abbildung 3.12: Funktion mittlere Bedienungsrate

→ μ ist die mittlere Bedienungsrate
→ $\frac{1}{\mu}$ ist die mittlere Bediennugsdauer (d.h. der Erwartungswert f_s definiertern Verteilung)
Die Anzahl $a(t)$ im Zeitintervall $[0, t)$ eingetroffener Kunden ist eine **Poisson-verteilte Zufallsgrösse** mit Erwartungswert (λt) es gilt

$$P[a(t) = k] = \frac{(\lambda t)^k}{k!} e^{-\lambda t} : k = 0, 1, 2 \tag{3.8}$$

- Ist das System im statistischen Gleichgewicht (stationär), so gilt dasselbe für die Anzahl Kunden b(t), die das System in diesem Zeitintervall verlassen.
- Im Folgenden wird die stationäre Verteilung der Anzahl $N(t)$ der zum Zeitpunkt t im System anwesenden Kunden hergeleitet (für den Fall, dass t gegen unendlich strebt und das System stationär ist)
- Beruht das Warteschlangenverhalten auf dem Markoveigenschaften, dann lässt sich das System als *single-server System mit unbegrenzter Kapazitätsauslastungsgrad und exponentialverteilten Interarrival- und Service-Zeitne* charakterisieren
- Es kann durch eine birth-death-Kette mit *birth und death Parametern* Ankunftsrate: $\lambda_i = \lambda$ for all i = 0,1,... und Service-Rate $\mu_i = \mu$ for all i = 1,2,..

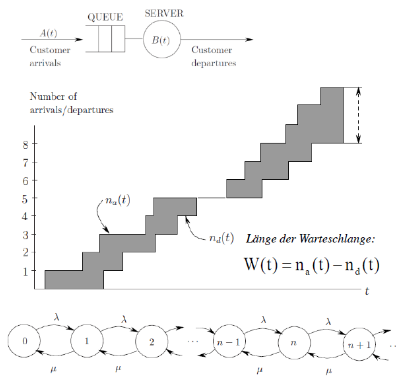


Abbildung 3.13: Zustandsübergangsraten Diagramm

→ Herleitung Chap03 ab Slide 56

Mittlere Anzahl Kunden im (stationären) System

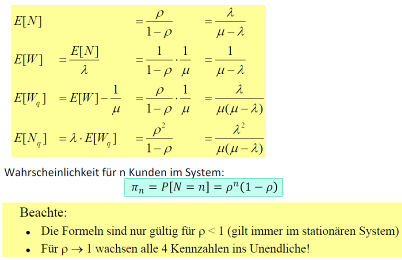
$$
\begin{aligned}
E[N] &= \sum_{i=0}^{\inf} i\pi_i = \sum_{i=0}^{\inf} i(1-p)p^1 \\
&= (1-p)\sum_{i=0}^{\inf} ip^i \\
&= (1-p)p\sum_{i=0}^{\inf} ip^{i-1} \\
&= \frac{p}{1-p} \\
\rightarrow &= \frac{\lambda}{\mu-\lambda}
\end{aligned} \tag{3.9}
$$

Wobei $\sum_{i=0}^{\inf} ip^{i-1} = \frac{1}{(1-p^2}$ die Ableitung von $\sum_{i=0}^{\inf} p^i = \frac{1}{1-p}$ ist

Mittlere Anzahl Kunden in der Schlange

$$
\begin{aligned}
E[N_q] &= \sum_{i=1}^{\inf}(i-1)\pi_i = \sum_{i=1}^{\inf} i\pi_i - \sum_{i=1}^{\inf}\pi_i = E[N]-(1-\pi_0) = \frac{p^2}{1-p} \\
E[N_q] &= \frac{p^2}{1-p} = \frac{\lambda^2}{\mu(\mu-\lambda)}
\end{aligned} \tag{3.10}
$$

3.4.1 Kennzahlen für Modell M/M/1



$$
\begin{aligned}
E[N] & & &= \frac{\rho}{1-\rho} &&= \frac{\lambda}{\mu-\lambda} \\
E[W] &= \frac{E[N]}{\lambda} & &= \frac{1}{1-\rho}\cdot\frac{1}{\mu} &&= \frac{1}{\mu-\lambda} \\
E[W_q] &= E[W]-\frac{1}{\mu} & &= \frac{\rho}{1-\rho}\cdot\frac{1}{\mu} &&= \frac{\lambda}{\mu(\mu-\lambda)} \\
E[N_q] &= \lambda\cdot E[W_q] & &= \frac{\rho^2}{1-\rho} &&= \frac{\lambda^2}{\mu(\mu-\lambda)}
\end{aligned}
$$

Wahrscheinlichkeit für n Kunden im System:

$$\pi_n = P[N=n] = \rho^n(1-\rho)$$

Beachte:

- Die Formeln sind nur gültig für $\rho < 1$ (gilt immer im stationären System)
- Für $\rho \rightarrow 1$ wachsen alle 4 Kennzahlen ins Unendliche!

Abbildung 3.14: Überblick der Kennzahlen für das Modell MM1

3.5 Das Model M/M/s

- Diese Schlange wird genauso wie das vorhergehende Modell M/M/1 analyisert
- Formel etwas kompliziert → Kein Problem, da Computer berechnet
- *Geburts- und Sterbeprozess* liegt zu Grunde mit ähnlichem Übergangsgraphen

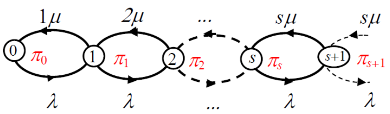


Abbildung 3.15: Abbildung eines Überblickgraphen für das Modell MMs

→ Analog zum M/M/1-Fall, eine zugehörige Übergangsmatrix und ein Gleichungssystem zur Bestimmung der stationären Zustandswahrscheinlichkeiten p_i i=0,1,2,... (Dies sei hier nicht explizit aufgeführt)

3.5.1 Lösung des Gleichungssystem

Die Lösung des Gleichungssystems ist:

$$\pi_i = \begin{cases} \frac{(s\rho)^i}{i!}\pi_0, & i = 1, \dots, s-1 \\ \frac{s^s \rho^i}{s!}\pi_0, & i = s, s+1, \dots \end{cases}$$

mit

$$\pi_0 = \left[1 + \frac{(s\rho)^s}{s!(1-\rho)} + \sum_{i=1}^{s-1} \frac{(s\rho)^i}{i!}\right]^{-1}$$

Zum Vergleich mit M/M/1:

$$\pi_i = \rho^i \pi_0$$
$$\pi_i = (1-\rho)\rho^i,$$
$$i = 0,1,2,\dots$$

mit

$$\pi_0 = \frac{1}{\sum_{i=0}^{\infty} \rho^i} = 1 - \rho,$$
$$i = 0,1,2,\dots$$

Abbildung 3.16: Lösung des Gleichungssystem für das Modell M/M/s

Daraus ergibt sich die *Wahrscheinlichkeit, dass ein ankommender Kunde warten muss :*

$$\zeta = P[\text{Anz. Kunden im System} \geq s] = \sum_{i=s}^{\infty} \pi_i =$$

$$= \pi_0 \frac{(s\rho)^s}{s!} \sum_{i=s}^{\infty} \rho^{i-s} = \pi_0 \frac{(s\rho)^s}{s!(1-\rho)}$$

Die *mittlere Anzahl Kunden im System* bzw. *in der Schlange* ist :

$$E[N] = \sum_{i=1}^{\infty} i\pi_i = s\rho + \frac{\rho\zeta}{1-\rho}$$

$$E[N_q] = \sum_{i=m+1}^{\infty} (i-m)\pi_i = \frac{\rho\zeta}{1-\rho}$$

Abbildung 3.17: Berechnungen für das Modell M/M/s

3.5.2 Verteilung der Wartezeit

Für stationäre M/M/s Schlange ist es möglich, eine explizite Form der Verteilung $F(x)$ der Wartezeit W_q abzuleiten:

$$F(x) = P[W_q \leq x] = 1 - \zeta e^{-s\mu(1-\rho)x}$$

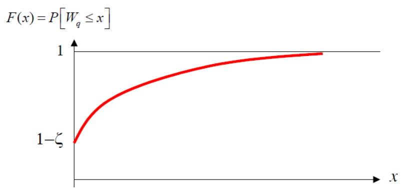


Abbildung 3.18: Verteilung der Wartezeit für das Modell M/M/s

Dabei ist ρ der Auslastungsgrad

$$\rho = \frac{\lambda}{s\mu}$$

und ζ die Wahrscheinlichkeit, dass der Kunde warten muss, d.h.

$$\zeta = \frac{(s\rho)^s}{s!(1-\rho)}\left[1 + \frac{(s\rho)^s}{s!(1-\rho)} + \sum_{i=1}^{s-1}\frac{(s\rho)^i}{i!}\right]^{-1}$$

Die Kenntnis der Wartezeitverteilung ist insbesondere dann von Interesse, wenn nicht nur die mittlere Wartezeit, sondern auch die *Wahrscheinlichkeit, dass der Kunde nicht mehr als eine vorgegebene Dauer wartet*, gefragt ist.

Abbildung 3.19: Verteilung der Wartezeit (II)für das Modell M/M/s

3.5.3 Kennzahlen für Modell M/M/s

$$
\begin{aligned}
E[N] & & &= \frac{2\rho}{1-\rho^2} \\
E[W] &= \frac{E[N]}{\lambda} & &= \frac{1}{1-\rho^2} \cdot \frac{1}{\mu} \\
E[W_q] &= E[W] - \frac{1}{\mu} & &= \frac{\rho^2}{1-\rho^2} \cdot \frac{1}{\mu} \\
E[N_q] &= \lambda \cdot E[W_q] & &= \frac{2\rho^3}{1-\rho^2}
\end{aligned}
$$

$$\pi_0 = P[N = 0] = \frac{1-\rho}{1+\rho}$$

$$\pi_n = P[N = n] = \begin{cases} 2\rho \frac{1-\rho}{1+\rho} : n = 1 \\ \frac{4\rho^n}{2} \frac{1-\rho}{1+\rho} : n = 2,3,\ldots \end{cases}$$

Beachte:

- $\rho = \frac{\lambda}{2\mu}$
- Die Formeln sind nur gültig für $\rho < 1$ (gilt immer im stationären System)
- Für $\rho \to 1$ wachsen alle 4 Kennzahlen ins Unendliche!

Abbildung 3.20: Überblick der Kennzahlen für das Modell MMs

Kapitel 4

Vorhersageverfahren

4.1 Vorhersagen im Operations Management

→ Auch Prognose oder Forecasting genannt
Dabei werden viele Entscheidungen des OM auf Vorhersagen getroffen:

- Wirksamkeit und Nutzen von Entscheidungen abhängig von ungewisser Zukunft
- Verwendung von Vorhersagen zur Evaluation alternativer Entscheidungen
- Planung und Vorbereitung von betrieblichen Abläufen

4.1.1 Unterscheidung bezüglich Entscheidungsebene (Tragweite)

- Strategisch (langfristig, $\geq$ 2 Jahre) → bspw. Entwicklung neuer Produkte, Investitionen in Anlagen, Standortentscheide
- Taktisch (mittelfrisitg, Monate bis 2 Jahre) → bspw. Produktionsprogramm-Planung, Budgetplanung, Einkauf
- Operativ (kurzfristig, Tage bis Monate) → Lagerhaltung, Ablaufplanung

weitere Unterscheidungskriterien:

- Ökonometrische Prognosen → Wirtschaftsindikatoren (Externe Quellen)
- Technologische Prognosen → Technologische Entwicklung bzgl Produkt, Analgen und Prozesse (interne und externe Quelle)
- Nachfrage-/ Bedarfsprognose → Bedarf von Endprodukte, Zwischenprodukte, Rohmaterialen, Analgenkapazität, Personal, Kapitel (interne und externe Quellen) → FOKUS

4.1.2 die 7 Schritte eines Forecasting

Bspw. von DisneyChap 04 Slide 7

1. Bestimmen der Verwendung der Vorhersage
2. Auswahl der zu prognostizierenden Elemente
3. Auswahl des Prognosemodells
4. Sammeln der erforderlichen Daten
5. Prognose erstellen
6. Ergebnis validieren und umsetzen

4.1.3 Nachfrage- Bedarfsprognose

Grundsätzlich unterscheidet man zwischen *unabhängiger* und *abhängiger* Bedarf

unabhängiger Bedarf: Bedarf an Endprodukten (bspw. Fahrräder)
abhängiger Bedarf: Bedarf der sich aus dem Bedarf von Endprodukten ergibt → Zwischenprodukte, Teilen, Rohmaterialen, Kapazität (bspw. Räder und Rahmen von Fahrräder)
⇒ Forecast bezieht sich auf **unabhängiger Bedarf** ⇒ unabhängiger Bedarf kann teilweise beeinflusst werden (bspw. Preisaktionen)

4.1.4 Die drei Grundgesetze des Forecastings

1. Prognosen sind immer falsch
2. Detaillierte Prognosen sind schlechter als aggregierte Prognosen

- aggregierte Prognosen haben kleinere Variabilität als detaillierte
- → Prinzip des **Variability Pooling**
- bspw. Prognosen für Produktefamilien besser als für einzelne Produkte

3. Je weiter Prognosen in die Zukunft reichen, desto weniger verlässlich sind sie

4.1.5 Die Kunst der Prognosen

- Forecasting ist sowohl Wissenschaft als auch 'Kunst'
- Unterstützung durch breites Spektrum von quantitativen Modellen
- Entscheidungsfindung häufig in Zusammenspiel mit **qualitative** Informationen und Intuition von Fachleuten
- Beizug von Spezialisten oft empfehlenswert

4.2 Prognoseverfahren: Eine Übersicht

4.2.1 Qualitative Verfahren

- Häufig in langfristigen Prognosen
- Entwickeln von Zukunftsszenarien mit Hilfe von Wissen, Erfahrung und Intuition von Experten (bspw. Komitte von Verantwortlichen aus Marketing, Produktion, Finanzen)
- **Delphi-Methode:**
  - Meist anonymisierte Befragung von Experten mittels Fragekatalog
  - Oft iterativer Prozess: Mehrere Befragungsrunden, bis Antworten 'stabilisiert'
  - Nach jeder Runde Aufarbeitung der Antworten und Entwurf eines neuen Fragekatalogs
  - bspw. Voraussichtlicher Zeitpunkt der Einführung einer neuen Technologie
- Marktstudien, Kundenbefragung
- Analogieschlüsse (bspw. Nachfrage eines Produkts schätzen durch vergleichbare Produkte)

4.2.2 Quantitative Verfahren

Zeitreihen-Modelle ('Time Series Models')

- Voraussage der zukünftigen Werten einer Variablen als eine Funktion von vergangenen Werten dieser Variablen → Prognose ausschliesslich gestützt auf einre Reihe vergangener Beobachtungen **Kein Erklärungsversuch** (bspw. Prognose der Nachfrage eines Produkts als Funktion der zurückliegenden Nachfrage)
- Diverse matehematische Modelle unterschliedlicher Komplexität: *Glättungsverfahren, Box-Jenkins-Verfahren, Shiskin Time Series, Kalman Filter*

Kausale Modelle

- Voraussage der zukünftigen Werte einer Variable ('interessierende Variable') als Funktion gewisser anderer Variablen ('ekrlärende Variablen') → Annahme einer kausalen Beziehung zwischen beiden Variablen **Erkärungsversuch** (bspw. Nachfrage eines Produkts in einer Region ('int. Variable') als Funktion vom Lohnniveau und Bevölkerungsdichte)
- Diverse mathematische Modelle: **Regressionsanalyse**

4.3 Komponenten einer Zeitreihe

Eine Zeitreihe ist eine *zeitliche Folge von beobachteten Werten ('Beobachtungen') einer gewissen Modellgrösse ('Variablen')*
Eine Zeitreihe X lässt sich in der Regel erklären durch die 'Überlagerung' von 4 Komponenten:

- Trend **T** → bspw. Nachfrage-Zuwachs eines Produkts
- Zyklische Schwankungen **C** → bspw. Konjunkturzyklen
- Saisonale Schwankungen **S** → bspw. Nachfrageunterschiede Sommer/Winter
- zufällige Schwankungen **I** → Autokorrelation

Überlagerung im Sinne einer *multiplikativen* oder *additiven* Verknüpfung

- **multiplikativ**: $X = T * C * S * I$
- **additiv**: $X = T + C + S + I$

4.4 Zeitreihen-Modelle: Glättungsverfahren

4.4.1 Bezeichnungen

i	Zeitperioden, $i = 1, 2, ...$
t	Zeitperiode der letzten Beobachtung*
	(d.h. für die Perioden $1, ..., t$ sind Beobachtungen verfügbar)
$f(t+\tau)$	Prognose für (zukünftige) Periode $t+\tau$, $\tau = 1, 2, ...$
$A(i)$	Beobachtung in Periode i, $i = 1, 2, ...$
$F(i)$	Geglätteter Schätzwert ("Smoothed Estimate") für Periode i, $i = 1, 2, ...$
$T(i)$	Geglätteter Trend ("Smoothed Trend") für Periode i, $i = 1, 2, ...$

Abbildung 4.1: Überblick der Bezeichnungen eines Glättungsverfahren

* Als Beobachtungswert sei im Folgenden der Bedarf eines Produkts angenommen.

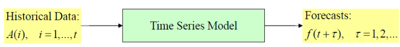


Abbildung 4.2: Abbildung eines Zeitreihen-Modells

4.4.2 Glättungsverfahren bei gleichmässigem Bedarf

Annahme: Bedarf unterliegt **nur zufälligen Schwankungen**, d.h. Zeitreihe hat nur Komponente I

Gleitender Durchschnitt ('Moving Average')

Parameter:
n: Anzahl berücksichtigte Perioden, $n \geq 1$
$F(t)$: Durchschnitt der letzten n Beobachtungen
Smoothed Estimate

$$F(t) = \frac{1}{n} \sum_{i=t-n+1}^{t} A(i) \tag{4.1}$$

Forecast

$$f(t+\rho) = F(t)\rho = 1, 2, .. \tag{4.2}$$

Dabei ist $F(t+1) = F(t) + \frac{A(t+1)}{n} - \frac{A(t-n+1)}{n}$

Gewichteter, gleitender Durchschnitt ('Weighted Moving Average')

Parameter:
n: Anzahl berücksichtigte Perioden, $n \geq 1$
w_i: Gewicht für Periode i, $i = 1, .., t$, wobei $w_i > 0$ und $\sum_{i=t-n+1}^{t} w_i = 1$
$F(t)$: Gewichteter Durchschnitt der letzten n Beobachtungen
Festlegen der Gewichte: Tendenziell den jüngeren Beobachtungen höhere Gewichte geben als den älteren
Smoothed Estimate

$$F(t) = \sum_{i=t-n+1}^{t} w_i * A(i) \tag{4.3}$$

Forecast

$$f(t+\rho) = F(t)\rho = 1, 2, .. \tag{4.4}$$

Exponentielle Glättung ('Exponential Smoothing')

Exponentielle Glättung ist einfach eine ausgefeilte **Gewichtete Moving-Average-Methode**, mit folgenden Merkregeln:

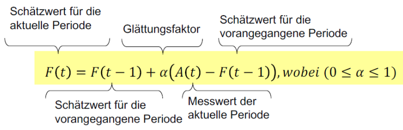


Abbildung 4.3: Merkregeln des Exponential Smoothing

Effekte des Exponential Smoothing:

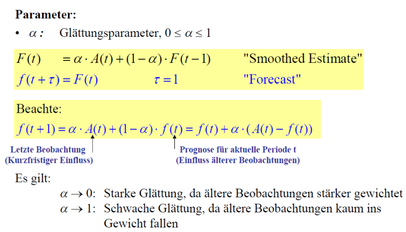


Abbildung 4.4: Effekt des Exponential Smoothing

→ Generelle Entwicklung Chap04 Slide 25

4.4.3 Glättungsverfahren bei einem Trend

Annahme: Bedarf unterliegt **zufälligen Schwankungen** und einem **Trend** → Zeitreihe umfasst die **zwei Komponenten I und T**

Exponentielle Glättung mit (linearem) Trend

Paraemeter:
α: Glättungsparameter für Schätzwert, $0 \leq \alpha \leq 1$
β: Glättungsparameter für Trend, $0 \leq \beta \leq 1$
Falls $\alpha, \beta \to 0$: starke Glättung,
falls $\alpha, \beta \to 1$: schwache Glättung
Smoothed Estimate

$$F(t) = \alpha * A(t) + (1 - \alpha) * [F(t-1) + T(t-1)] \tag{4.5}$$

Smoothed Trend

$$T(t) = \beta * [F(t) - F(t-1)] + (1 - \beta) * T(t-1) \tag{4.6}$$

Forecast

$$f(t + \rho) = F(t) + \rho * T(t)\rho = 1, 2, .. \tag{4.7}$$

Aufbau

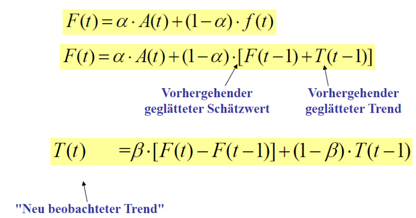


Abbildung 4.5: Aufbau der Exponentielle Glättung mit Trend

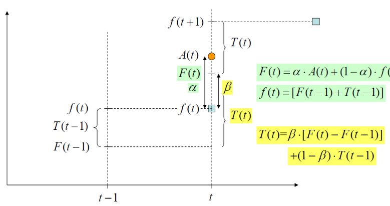


Abbildung 4.6: Illustration der Berechnung der Exponentielle Glättung mit Trend

Glättungsverfahren bei Trend und Saisonalität

D.h. Zeitreihe umfastt die zwei Komponenten I und T und S.
Winters Methode
Parameter:
α: Glättungsparameter für Schätzwert $0 \leq \alpha \leq 1$
β: Glättungsparameter für Trend $0 \leq \beta \leq 1$
γ: Glättungsparameter für Saisonalität $0 \leq \gamma \leq 1$
Falls $\alpha\beta\gamma \to 0$: Starke Glättung
Falls $\alpha\beta\gamma \to 1$: Schwache Glättung *Smoothed Estimate*

$$F(t) = \alpha * \frac{A(t)}{c(t-N)} + (1 - \alpha) * [F(t-1) + T(t-1)] \tag{4.8}$$

Smoothed Trend

$$T(t) = \beta * [F(t) - F(t-1)] + (1 - \beta) * T(t-1) \tag{4.9}$$

Smoothed Saisonality

$$c(t) = \gamma * \frac{A(t)}{F(t)} + (1-\gamma) * c(t-N) \tag{4.10}$$

Forecast

$$f(t+\rho) = [F(t) + \rho * T(t)] * c(t+\rho-N), t+\rho = N+1, ..., 2N \tag{4.11}$$

Initialisierung c(t) mit Saisonfaktor

$$c(t) = \frac{A(t)}{\sum_{i=t-N+1}^{t} \frac{A(t)}{N}} \tag{4.12}$$

4.4.4 Beurteilung der Qualität von Prognosenverfahren

Auflistung der häufig benutzen Masse zur Qualitätsbeurteilung.

Mittlere Absolute Abweichung (MAD, 'Mean Absolute Deviation')

$$MAD = \frac{\sum_{i=1}^{m} |f(i) - A(i)|}{m} \tag{4.13}$$

Mittlere Quadratische Abweichung (MSD, 'Mean Square Deviation')

$$MSD = \frac{\sum_{i=1}^{m} [f(i) - A(i)]^2}{m} \tag{4.14}$$

Verzerrung ('BIAS')

$$BIAS = \frac{\sum_{i=1}^{m} f(i) - A(i)}{m} \tag{4.15}$$

Verwendung der Qualitätsmasse (in konkreter Anwendung)

Unter Ausnützung von (möglichst umfangreichen) Vergangenheitsdaten:

- Finden von geeigneten **Parametereinstellungen** ('Kalibrierung)
- **Vergleich** verschiedener **Prognoseverfahren**

4.5 Kausale Modelle: Lineare Regression und Trendprojektion

4.5.1 Einfache lineare Regression

- **Zweck**: Quantifizierung des funktionalen Zusammenhangs zwischen X (unabhängige Variable) und Y (abhängige Variable)
- **Hypothese**: Linearer Zusammenhang $Y = A + BX +$ Zufallsvariable mit E-Wert 0 $\rightarrow$ *beachte*: Auch andere Beziehungen auf lin. Fall zurückführbar
- Beste Schätzungen für $\hat{a}$ und $\hat{b}$ von A und B anhang von Beobachtungen finden
- X: deterministische Variable mit Werten $x_1, ..., x_n$
- Y: Zufallsvariable
- $x := \frac{1}{n} \sum_{j=1}^{n} x_j$, $y := \frac{1}{n} \sum_{j=1}^{n} y_j$
- Beste Schätzungen für $\hat{a}$ und $\hat{b}$: Wähle $\hat{a}$ und $\hat{b}$ so, dass $Q = \sum_{j=1}^{n} [y_j - E(Y_j)]^2$ minimal ist

$$Q = \sum_{j=1}^{n} [y_j - E(Y_j)]^2 = \sum_{j=1}^{n} [y_j - A - Bx_j]^2$$

$$\frac{\partial Q}{\partial A}(\hat{a}, \hat{b}) = -2 \sum_{j=1}^{n} [y_j - \hat{a} - \hat{b}x_j] = 0$$

$$\hat{a} = \bar{y} - \hat{b}\bar{x}$$

$$\frac{\partial Q}{\partial B}(\hat{a}, \hat{b}) = -2 \sum_{j=1}^{n} [y_j - \hat{a} - \hat{b}x_j] \cdot x_j = 0$$

$$\ldots. \quad \hat{b} = \frac{\sum y_j x_j - n\bar{y}\,\bar{x}}{\sum x_j^2 - n\bar{x}^2}$$

Abbildung 4.7: Berechnung für linReg

4.5.2 Trendprojektion

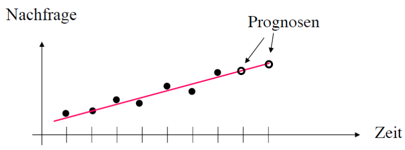


Abbildung 4.8: Abbildung einer Trendprojektion

Beispiel:

Jahr	xj	Nachfrage yj		xj²	xj.yj		Yj=a+b*xj
1997	1	1000		1	1000		1129.1
1998	2	1300		4	2600		1344.8
1999	3	1800		9	5400		1560.6
2000	4	2000		16	8000		1776.4
2001	5	2000		25	10000		1992.1
2002	6	2000		36	12000		2207.9
2003	7	2200		49	15400		2423.6
2004	8	2600		64	20800		2639.4
2005	9	2900		81	26100		2855.2
2006	10	3200		100	32000		3070.9
2007	11			121			3286.7
2008	12			144			3502.4
Summe	55	21000		385	133300		
Mittelwert	5.5	2100					
b =	215.76		a =	913.33			

$$\hat{b} = \frac{\sum y_j x_j - n\bar{y}\,\bar{x}}{\sum x_j^2 - n\bar{x}^2} \qquad \hat{a} = \bar{y} - \hat{b}\bar{x}$$

Abbildung 4.9: Beispiel einer Trendprojektion

Prognose für 2007 auf Quartale verfeinern

Basis: Jahre 2004, 2005, 2006: Quartaldaten für die Nachfrage

1) Bestimmung der Saisonfaktoren

	1. Quart.	2. Quart.	3. Quart.	4. Quart.	Total Jahr	
2004	520	730	820	530	2600	
2005	590	810	900	600	2900	
2006	650	900	1000	650	3200	
	1760	2440	2720	1780	8700	Total über die 3 Jahre
	586.7	813.3	906.7	593.3	725.0	Quartalmittel
	0.809	**1.122**	**1.251**	**0.818**		**Saisonfaktoren**

593.3/725.0

2) Saisonbereinigung der Daten

	1. Quart.	2. Quart.	3. Quart.	4. Quart.
2004	642.6	650.7	655.7	647.6
2005	729.1	722.0	719.7	733.1
2006	803.3	802.3	799.6	794.2

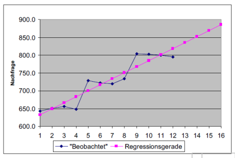
650/0.818

Abbildung 4.10: Beispiel einer Trendprojektion mit saisonalen Schwankungen I

3) Lineare Regression mit 12 Beobachtungen: Saisonbereinigte Prognosen für 2007

	xj	yj		Regressionsgerade
2004 1Q	1	642.6		632.3
2004 2Q	2	650.7		649.2
2004 3Q	3	655.7		666.1
2004 4Q	4	647.6		683.0
2005 1Q	5	729.1		699.9
2005 2Q	6	722.0		716.8
2005 3Q	7	719.7		733.7
2005 4Q	8	733.1		750.6
2006 1Q	9	803.3		767.5
2006 2Q	10	802.3		784.4
2006 3Q	11	799.6		801.3
2006 4Q	12	794.2		818.2
2007 1Q	13			**835.1**
2007 2Q	14			**852.0**
2007 3Q	15			**868.9**
2007 4Q	16			**885.8**
b =	**16.9**		**a =**	**615.4**

Abbildung 4.11: Beispiel einer Trendprojektion mit saisonalen Schwankungen II



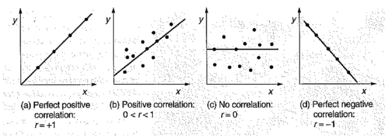
4) Saisonberücksichtigende Prognosen für 2007

	Saison-faktoren	Saison-bereinigt	Saison-berücksichtigend
2007 1Q	0.809	835.10	**676**
2007 2Q	1.122	852.00	**956**
2007 3Q	1.251	868.90	**1087**
2007 4Q	0.818	885.80	**725**

Abbildung 4.12: Beispiel einer Trendprojektion mit saisonalen Schwankungen III

4.5.3 Korrelationskoeffizienten für lineare Regressionen

- **Zweck:** Korrelationskoeffizienten geben die Stärke der linearen Abhängigkeit wieder
- Der Korrelationskoeffizienten r ist ein Wert zwischen -1 (negative Werte: antikorreliert) und +1 (positive Werte: korreliert)
- Werte um 0 lassen vermuten, dass keine Korrelation vorliegt. Die Berechnung ist ähnlich wie die für die lineare Regression.



Korrelationskoeffizient

$$r_{xy} := \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2} \cdot \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

Alternative Berechnung von r_{xy}:

$$r_{xy} := \frac{n\sum x_j y_j - \sum x_j \sum y_j}{\sqrt{\left[n\sum x_j^2 - \left(\sum x_j\right)^2\right]\left[n\sum y_j^2 - \left(\sum y_j\right)^2\right]}}$$

Abbildung 4.13: Korrelationskoeffizienten

Pearson'sche Korrelationskoeffizienten

obere Zeile: Wiedergabe der Streuung der Punktewolke sowie die generelle Richtung der linearen ABhängig von x und y
Mittlere Zeile: Keine Information zur Steilheit. Verläuft die Punktewolke exakt waagerecht (mittleres Bild) kann aufgrund von $Var(Y) = 0$ gar kein Korrelationskoeffizienten berechnet werden
Untere Zeile: Schwachpunkt, nichtlineare Abhängigkeiten nur unzureichend erfasst

Weitere Techniken: → Chap04 Slide 62

4.6 Monitoring und Steuerung von Forecasts

4.6.1 Tracking Signal

- **Zweck**: Rollendes (nachgeführtes quantitatives Mass dafür, wie gut ein Forecast aktuelle Werte vorhersagt)
- **Berechnung**: Verhältnis der laufenden Summe der Forecast-Fehler (*Running sum of the forecast errors [**RFSE**]*) zum mittleren absoluten Fehler (*Mean absolute deviation [**MAD**]*)

Tracking Signal:

$$\frac{RFSE}{MAD} = \frac{\sum(A_i - f_i)}{\frac{\sum |A_i - f_i|}{n}} \tag{4.16}$$

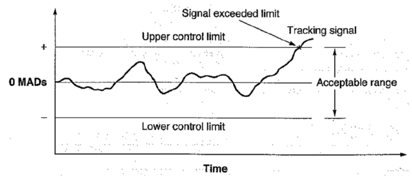


Abbildung 4.14: Abbildung der Grafik zum Tracking Signal

Kapitel 5

Mathematische Optimierung

5.1 Optimierungsmethoden

5.1.1 Problemformulierung

Ein allgemeines Optimierungsproblem (bzw. Optimierungsmodell) $\prod$ kann wie folgt geschrieben werden: $\prod$: max$f(x)$: $x \in S$, wobei $S \subseteq \mathbb{R}^n$

- Ein Problem heisst **eingeschränkt, falls S** $\subset \mathbb{R}^n$ $\rightarrow$ enthalten Restriktionen (Zulässigkeitsbedingungen), welche die Menge der **zulässigen Lösungen S** $\subseteq \mathbb{R}^n$ auf eine echte Teilmenge von $\mathbb{R}^n$ beschränken $\Rightarrow$ Aufgabe eine geeignete funktionale Beschreibung der Restriktionen zu finden
- Ein Problem heisst **uneingeschränkt, falls** $S = \mathbb{R}^n$
- Allgemein lässt sich sage, dass eingeschränkte Probleme typischerweise schwieriger zu lösen sind als uneingeschränkte

5.1.2 Eingeschränkte Optimierungsprobleme

(i) Funktionale Restriktionen:
Der Lösungsraum S wird durch

- eine Anzahl $p \geq 0$ von Ungleichungen der Form $g_i(x) \leq 0$ und
- eine Anzahl $q \geq 0$ von Gleichungen der Form $h_j(x) = 0$ definiert: S = $x \in G_1$ x G_2 x.. x G_n : $g_i(x) \leq 0, h_j(x) = 0$, i=1..p

(ii) Nicht-funktionale Restriktionen:
Der Lösungsraum S ist durhc gewisse 'nichtfunktionale' Restriktionen definiert, bspw. logische Prädikate, welche die zulässigen Elemente aufgrund bestimmter Eigenschaften aus einer Grundmenge auswählen:
$\hat{S} = x \in G_1$ x G_2 x .. x G_n : 'x hat gewisse Eigenschaften $\rightarrow$ bspw. $x \in \mathbb{Z}^n$: x ist eine Permutation der Zahlen 1,..,n

5.1.3 Maximierungs- und Minimierungsprobleme

Optimierungsprobleme können Maximierungs- oder Minimierungsprobleme sein. Es gilt:

$$max\{f(x) : x \in S\} = -min\{-f(x) : x \in S\} \tag{5.1}$$

und $x^* \in S$ ist genau dann eine Optimallösung des Maximierungsproblems, wenn x^* eine Optimallösung des Minimierungsproblems ist.

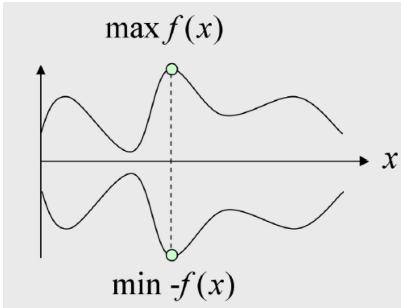


Abbildung 5.1: Abbildung der Min-Max-Optimierung

5.1.4 Nachbarschaftsbegriffe

Eine auf S definierte nachbarschaft ist eine (mengenwertige) Funktion der Form $N : S \rightarrow P(S)$ welche jedem Punkt $x \in S$ eine Menge von benachbarten Punkte N(x) zuordnet, wobei gelten soll, dass $x \in N(x)$.
$N(x)$ wird als die Nachbarschaft von x bezeichnet.

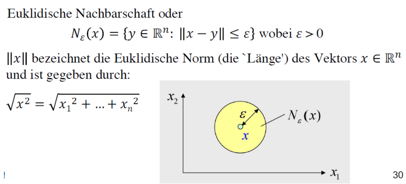


Abbildung 5.2: Abbildung der euklidische Nachbarschaft

5.1.5 globales und lokales Optimum

Ein $x^* \in S$ ist (bzgl. Nachbarschaft) eine lokale Optimallösung und $f(x^*)$ ein lokales Optimum von $\prod$ falls $\epsilon > 0$ existiert, so dass $f(x^*) \geq f(x)$ für alle $x \in N_\epsilon(x^*) \cap S$

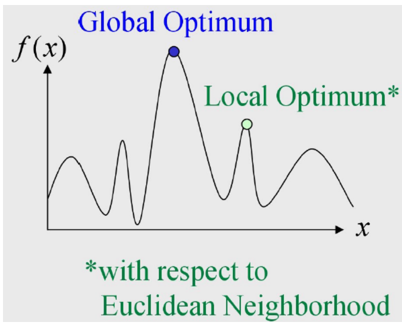


Abbildung 5.3: Funktion globales bzw lokales Optimum

5.1.6 Konvexe Menge

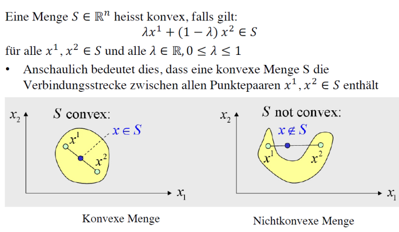


Abbildung 5.4: Konvexe Menge

5.1.7 konvexe und konkave Funktion

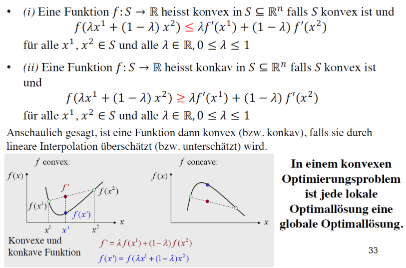


Abbildung 5.5: konvexe vs konkave Funktion

5.1.8 Konvex vs. nicht-konvex

Konvexe Optimierung beschäftigt sich mit der Maximierung (bzw. Minimierung) einer konkaven (bzw. konvexen) Zielfunktion f über einem konvexen Lösungsraum S.
Sie spielen eine herausragende Rolle in der Optimierungstheorie, da sich die Suche nach einem globalen Optimum auf die Suche nach einem lokalen Optimum reduziert.

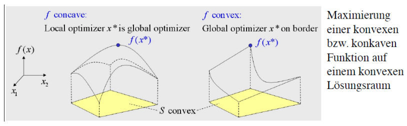


Abbildung 5.6: Nicht konvex

5.1.9 global vs. lokal

Unter globaler Optimierung wird die Aufgabe verstanden, eine globale Optimallösung für ein Optimierungsproblem zu finden (oder festzustellen, dass keine solche existiert).
Bei lokaler Optimierung begnügt man sich mit dem Auffinden einer lokal optimalen Lösung. Trivialerweise ist das erstere mindestens so schwierig wie das zweite. Viele Bereiche der Optimierungstheorie ausserhalb der konvexen Optimierung befassen sich hauptsächlich mit lokaler Optimierung, wie beispielsweise das Gebiet der allgemeinen nicht-linearen Optimierung.

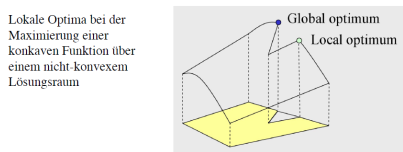


Abbildung 5.7: global vs local

5.1.10 Kombinatorische Nachbarschaft

Eine diskrete ('kombinatorische') Nachabrschaft in $\mathbb{Z}^n$ ist gegeben durch $N(x) = \{x' \in \mathbb{Z}^n : x'$

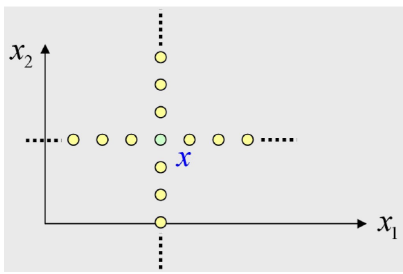
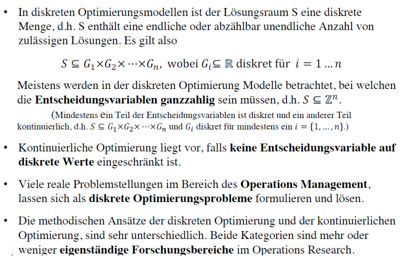


Abbildung 5.8: Beispiel einer kombinatorische Nachbarschaft für $\mathbb{R}^2$

5.1.11 diskret vs. kontinuierlich



- In diskreten Optimierungsmodellen ist der Lösungsraum S eine diskrete Menge, d.h. S enthält eine endliche oder abzählbar unendliche Anzahl von zulässigen Lösungen. Es gilt also

$$S \subseteq G_1 \times G_2 \times \cdots \times G_n,\ \text{wobei } G_i \subseteq \mathbb{R} \text{ diskret für } i = 1 \ldots n$$

Meistens werden in der diskreten Optimierung Modelle betrachtet, bei welchen die **Entscheidungsvariablen ganzzahlig** sein müssen, d.h. $S \subseteq \mathbb{Z}^n$.

(Mindestens ein Teil der Entscheidungsvariablen ist diskret und ein anderer Teil kontinuierlich, d.h. $S \subseteq G_1 \times G_2 \times \cdots \times G_n$ und G_i diskret für mindestens ein $i = \{1, \ldots, n\}$.)

- Kontinuierliche Optimierung liegt vor, falls **keine Entscheidungsvariable auf diskrete Werte** eingeschränkt ist.
- Viele reale Problemstellungen im Bereich des **Operations Management**, lassen sich als **diskrete Optimierungsprobleme** formulieren und lösen.
- Die methodischen Ansätze der diskreten Optimierung und der kontinuierlichen Optimierung, sind sehr unterschiedlich. Beide Kategorien sind mehr oder weniger **eigenständige Forschungsbereiche** im Operations Research.

Abbildung 5.9: diskret vs kontinuierlich

5.1.12 linear vs. nichtlinear

- lineare Optimierung beschäftigt sich mit der Optimierung einer linearen Zielfunktin unter Berücksichtigung einer Anzahl linearer Restriktionen (Ungleichungen oder Gleichungen)
- Lineare Modelle liegen am Ursprung des Operations Research und weisen tief reichende mathematische Strukturen auf, welche es heutzutage erlauben, auch grosse Probleminstanzen effizient zu lösen
- Von zentraler Bedeutung für die Praxis sind Modelle der ganzzahligen Linearen Programmierung, bei welchen alle (oder einige) Entscheidungsvariablen auf ganzzahlige Werte beschränkt (also diskret) sind.

5.1.13 Niveaumengen

Niveaumengen spielen insbesondere bei der graphischen Darstellung von zwei- und dreidimensionalen Funktionen eine Rolle.

- Sei $f : S \rightarrow \mathbb{R}$ mit $S \subseteq \mathbb{R}^n$. Die Niveaumenge L_a von f zum Niveau $\alpha \in \mathbb{R}$ ist gegeben durch $L_\alpha(x) = \{x \in S : f(x) = \alpha$
- Für eine n-dimensionale Funktion f sind die Niveaumengen von der 'Dimension' n-1 (oder kleiner)
- Im Fall n=2 sind die Niveaumengen typischerweise 'Linien' ($\rightarrow$ 1D) und werden auch Niveaulinien (geografisch: Höhenkurven) genannt
- Falls $f : \mathbb{R}^n \rightarrow \mathbb{R}$ eine lineare Funktion ist, d.h. $f(x) = c_1x_1 + ..c_nx_n$ für ein $c \in \mathbb{R}^n$, entsprehceh die Niveaumengen parallelen Geraden, wobei der Vektor c orthogonal zu diesen Geraden ist und das Niveau α zunimmt, wenn die Geraden in Richtung von a 'parallel verschoben' werden.
- Für n=2 sind sie typischerweise 'Flächen' (2D Mengen) und werden auch Niveauflächen genannt

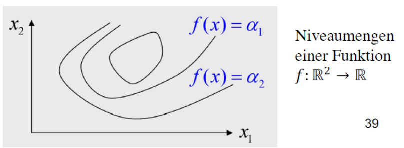


Abbildung 5.10: Abbildung von Niveaumengen

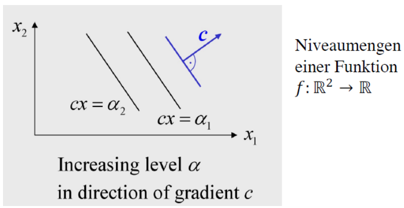


Abbildung 5.11: Abbildung von Niveaumengen II

5.1.14 Halbräume und Hyperebenen

Sei $a \in \mathbb{R}^n$ und $b \in \mathbb{R}$. Die Mengen $H = \{x \in \mathbb{R}^n : ax \leq b\}$ wird ein **Halbraum** und die Menge $H' = \{x \in \mathbb{R}^n : ax = b\}$ eine **Hyperebene** in $\mathbb{R}^n$ genannt. H' heisst die definierende Hyperebene von H.

Hyperebene

Für eine Hyperebene $H' = \{x \in \mathbb{R}^n : ax = b\}$ wird der Vektor a als Normalenvektor bezeichnet.

→ Es gilt: $a(x - x_0) = ax - ax_0 = b - b = 0$ für alle $x \in H'$, d.h. a ist orthogonal zu allen Parallelvektoren von H'

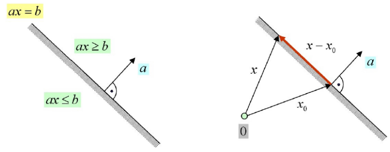


Abbildung 5.12: Abbildung einer Hyperebene

5.1.15 Polyeder und Polytop

Polyeder

Ein **Polyeder P** in $\mathbb{R}^n$ ist der Durchschnitt einer endlichen Anzahl von Halbräumen in $\mathbb{R}^n$ d.h. $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ für ein $A \in \mathbb{R}^{m \times n}$ und $b \in \mathbb{R}^m$. Die Mengen $\{x \in \mathbb{R}^n : a^i x = b_i$, heissen die definierenden Hyperebenen von P.
⇒ Ein Polyeder ist also die Lösungsmenge eines linearen Ungleichungssystems.

Polytop

Ein **Polytop** in $\mathbb{R}^n$ ist ein begrenztes Polyeder, d.h. ein Polyeder $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ für welches $l, u \in \mathbb{R}^n$ existieren, so dass P $=$ $x \in \mathbb{R}^n$: $Ax \leq b$, $l \leq x \leq u$

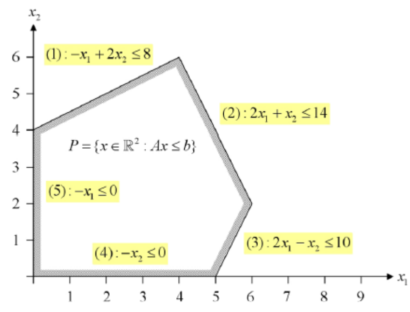


Abbildung 5.13: Beispiel eines Polytops in $\mathbb{R}^2$

5.1.16 Geometrische Lösung von Linearen Programme in $\mathbb{R}^2$ und $\mathbb{R}^3$

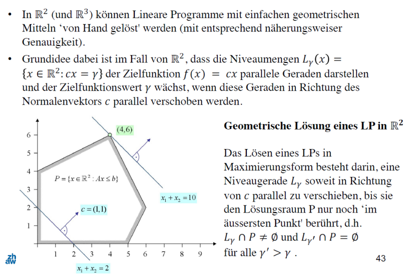


Abbildung 5.14: Geometrische Lösung von Linearen Programme

5.1.17 Simplex-Algorithmus

Input: Eine instanz eines Linearen Programms $\prod$:
max cx: $x \in \mathbb{R}^n$ mit $P = x \in \mathbb{R}^n$: $Ax \leq b$ $\;A \in \mathbb{R}^{mXn}$ und rank(A) = n. Dazu eine zulässige Basisauswahl $B \subseteq \{1, .., m\}$, $|B| = n$

Output:
Eine optimale Eckpunktlösung $v \in P$, falls $\prod$ beschränkt ist. Andernfalls eine 'Richtung' $d \in \mathbb{R}^n$, entlang welcher die Zielfunktion unbeschränkt wachsen kann.

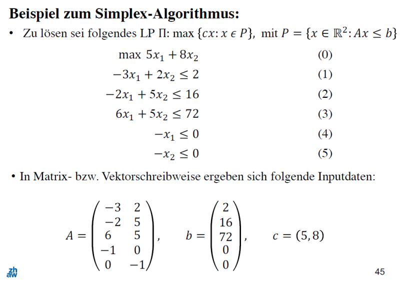


Abbildung 5.15: Beispiel zum Simplex-Algorithmus I

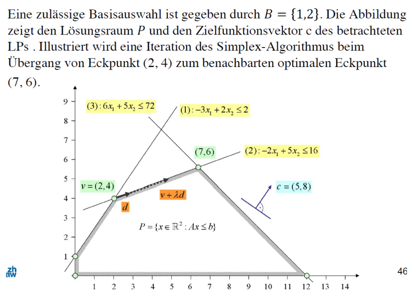


Abbildung 5.16: Beispiel zum Simplex-Algorithmus II

5.1.18 ganzzahlig-lineare Programmierung (engl. Integer Linear Program)

(i) LP: Der Lösungsraum eines Linearen Programms entspricht einem Polyeder, und ein LP kann allgemein geschrieben werden als: $LP : max\{cx : x \in P\}$ mit $P = \{x \in \mathbb{R}^n : Ax \leq b\}$
(ii) ILP: Ein ganzzahlig-lineares Programm hat somit als Lösungsraum die ganzzahligen Vektoren eines Polyeder und kann geschrieben wreden als: $ILP : max\{cx : x \in P \cap \mathbb{Z}^n\}$ mit $P = \{x \in \mathbb{R}^n : Ax \leq b\}$

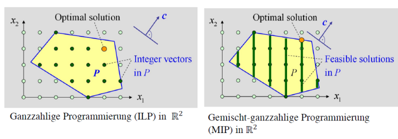
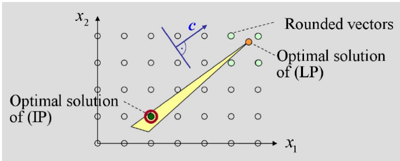


Abbildung 5.17: ILP

Die meinung, dass eine nahezu optimale Lösung für ein ILP einfach gefunden werden kann, indem das entsprechende LP (ohne Ganzzahligskeitsforderung) gelöst wird und die gefundene Optimallösung auf- bzw. abgerundet wird, ist weit verbreitet
Dabei kann die Optimallösung des ILP beliebig weit von der Optimallösung des LP 'entfernt' liegen. ⇒ Bei dieser Überlegung wird übersehen, dass durch Runden eventuell gar keine zulässige Lösung des ILP konstruiert werden kann.



Runden ist kein Lösungsansatz für ILP

Abbildung 5.18: ILP II

5.2 CMPL

<Coliop|Coin> Mathematical Programming Language ist eine mathematische Programmiersprache und ein System für mathematische Programmierung und Optimierungen für lineare Optimierungsprobleme.