

数字控制系统

贺允东 编著

人民邮电出版社

数 字 控 制 系 统

贺允东 编著

人民邮电出版社

内 容 提 要

本书主要介绍了数字控制系统中的基本理论，数字控制系统的分析和综合，以及如何将微型计算机用于动态控制系统，以实现实时控制。内容丰富，分析详尽。为了帮助读者学习，书中有较多数量的例题和习题，并在附录中补充了一些数学工具和基础知识。

本书主要对象是控制专业的研究生和高年级大学生，也可供有关工程技术人员以及高等学校其它专业的学生和教师参考。

数 字 控 制 系 统

贺允东 编著

责任编辑：梁 凝

*
人民邮电出版社出版
北京东长安街27号

北京印刷一厂印刷
新华书店北京发行所发行
各地新华书店经售

*
开本：787×1092 1/16 1986年10月 第一版
印张：29 页数：232 1986年10月北京第一次印刷
字数：722千字 印数：1—5,500册
统一书号：15045·总3249—有5474
定价：5.75元

前　　言

控制理论和系统在工业、农业、交通运输、军事和经济管理等部门的重要作用，已是人所共知的事实。

由于数字系统具有稳定性好、精度高和运算速度快等优点，加以数字计算机特别是微型计算机的迅速发展，数字控制系统在现代科学技术中的地位愈益重要，不仅一些模拟控制系统已经被它所代替，而且它已经渗透到飞机的自动驾驶、宇宙运载工具的轨道控制、通信的自动化、交通的自动调度指挥和机器人的操纵等领域。例如，将数字控制系统用于生产过程，不仅可以节省人力，还能提高产量和确保品质，并对事故进行预警监视；在通信中应用数字控制，可以做到自动管理，保证通信畅通；采用微处理器实现数字控制的雷达，可做到自动搜索和定位测距。

基于上述原因，在邮电出版社的大力支持下，我于 1983 年下半年开始编写此书。为了能将数字控制系统付诸实用，除去要掌握模拟控制系统的基本理论外，重要的是，要了解数字控制系统本身的特点，以及如何将数字控制系统同数字计算机结合起来。因此，本书的重点是研究数字系统的理论，数字控制系统的分析和综合，以及微型计算机在数字控制系统中的应用。

第一章概略介绍数字控制系统的框图及其优点。第二章讨论信号的变换，数字信号和量化，以及模拟信号的抽样。第三章介绍数字系统的一个重要数学工具—— z 变换。除去 z 变换的定义和性质外，着重讨论 z 传递函数，约量抽样，变形 z 变换，多速率抽样系统的 z 变换，以及数字系统的信号流图。第四章状态变量分析法，重点讨论数字系统和抽样数据系统的状态方程的建立及其解，状态转移矩阵的计算，抽样数据系统的状态图，以及非均匀抽样数据系统的状态变量分析法。第五章讨论在现代控制理论中占有重要地位的能控性和能观测性。第六章研究数字系统的稳定性，以及鞠利稳定判据、双线性变换和李亚普诺夫稳定判据。第七章时域分析，讨论抽样周期对输出波形和系统稳定性的影响，抽样数据系统的阶跃响应， z 平面上的根轨迹，稳态误差，以及数字仿真。第八章讨论抽样数据系统的两种频域分析法：伯德图，以及用 z 变换和无穷级数法的奈魁斯特判据。第九、十章讨论数字控制系统的设计。第九章用基于频域图解法的经典法设计数字控制系统，主要内容有：用无源网络实现数字控制器，直接设计法，以及用数字计算机实现数字控制器。最后，介绍最快响应和无振荡数字控制系统的设计。第十章讨论以状态变量为基础的数字控制系统的设计，主要内容有：极点配置，求反馈矩阵 G ，状态观测器，以及组合系统的动态特性。第十一章讨论以线性二次型最优控制为中心的最优控制问题，着重介绍动态规划和庞特里雅金最小值原理。第十二章讨论微处理器在数字控制系统中的应用，主要内容有：截尾误差和舍入误差，极限环，以及用微处理器的控制系统。

本书是在 1983 年北方交通大学研究生使用的同名讲义的基础上编写的，可作为研究生和高年级大学生的教材或教学参考书，也可供工程技术人员自学。阅读本书，应该具有线性系统、模拟控制系统和数字计算机等方面的基本知识。为了帮助不完全具备上述知识的读者阅读本书，在附录中补充了一些必要的数学工具和其它基础知识。

本书突出了基本原理和基本概念，避免过多的数学推导；力求叙述简练，理论同实际结合；还融合了作者的研究成果。为便于读者学习，书内有较多数量的例题和习题。

兰州铁道学院班冀超教授、北方交通大学叶杭教授和潘恂如副教授在百忙中仔细审阅了全部手稿，感谢他们提出了许多宝贵意见；感谢汪庆萱同志为本书绘制了全部制版图。

由于编写时间紧迫和作者学识有限，本书定有不妥之处，诚恳地希望得到批评指正。

贺允东

1985年3月

于北京 北方交通大学

目 录

| | |
|---|----|
| 第一章 引论 | 1 |
| 第一节 概述..... | 1 |
| 第二节 数字控制系统的优点..... | 1 |
| 第二章 信号的变换 | 3 |
| 第一节 概述..... | 3 |
| 第二节 数字信号和量化..... | 3 |
| 一、数字信号..... | 3 |
| 二、量化..... | 6 |
| 第三节 D/A 和 A/D 变换器..... | 8 |
| 一、D/A 变换器 | 8 |
| 二、A/D 变换器 | 9 |
| 第四节 模拟信号的抽样..... | 11 |
| 一、抽样过程..... | 11 |
| 二、抽样定理..... | 14 |
| 三、理想抽样..... | 15 |
| 四、 $F^*(s)$ 的特性 | 19 |
| 第五节 保持器..... | 20 |
| 附录..... | 25 |
| 2.1 拉普拉斯变换 | 25 |
| 2.2 确定 ξ 平面的积分路径 | 28 |
| 2.3 求 $\frac{1}{2\pi j} \int_{L_1} \frac{F(\xi)}{1-e^{-(s-\xi)T}} d\xi$ 和 $\frac{1}{2\pi j} \int_{L_2} \frac{F(\xi)}{1-e^{-(s-\xi)T}} d\xi$ | 29 |
| 习题..... | 30 |
| 第三章 z 变换 | 33 |
| 第一节 概述..... | 33 |
| 第二节 z 变换的定义 | 33 |
| 第三节 z 反变换 | 37 |
| 第四节 z 变换的性质和定理 | 41 |
| 第五节 s 平面和 z 平面间的映射 | 48 |
| 一、 $F(z)$ 、 $F^*(s)$ 和 $F^*(j\omega)$ 间的关系 | 48 |
| 二、s 平面和 z 平面间的映射关系 | 49 |
| 第六节 脉冲传递函数和 z 传递函数..... | 51 |
| 一、脉冲传递函数和 z 传递函数 | 51 |
| 二、接有零阶保持器的系统..... | 54 |
| 第七节 抽样时刻间的响应..... | 56 |

| | |
|---------------------------------|------------|
| 一、约量抽样..... | 56 |
| 二、变形 z 变换..... | 59 |
| 第八节 抽样数据系统的信号流图..... | 63 |
| 一、抽样信号流图法..... | 63 |
| 二、直接信号流图法..... | 68 |
| 第九节 多速率抽样数据系统..... | 72 |
| 一、慢-快多速率抽样系统 | 73 |
| 二、快-慢多速率抽样系统 | 75 |
| 三、多速率全数字系统..... | 77 |
| 四、多速率抽样系统中插入低速抽样器..... | 78 |
| 五、闭环多速率抽样系统..... | 80 |
| 附录..... | 83 |
| 3.1 拉氏变换和 z 变换表 | 83 |
| 3.2 二阶连续系统的传递函数 | 84 |
| 习题..... | 85 |
| 第四章 状态变量分析法..... | 92 |
| 第一节 概述..... | 92 |
| 第二节 连续系统的状态方程..... | 92 |
| 一、线性时不变系统..... | 94 |
| 二、线性时变系统..... | 97 |
| 第三节 数字系统的状态方程..... | 100 |
| 一、抽样数据系统的状态方程..... | 100 |
| 二、线性离散系统的状态方程..... | 102 |
| 第四节 数字系统状态方程的解..... | 102 |
| 一、抽样数据系统状态方程的解..... | 102 |
| 二、线性离散系统状态方程的解..... | 106 |
| 第五节 状态变量表达式和传递函数的关系..... | 107 |
| 第六节 特征方程, 特征值和特征向量..... | 110 |
| 一、定义..... | 110 |
| 二、特征值的性质..... | 111 |
| 三、特征向量的性质..... | 112 |
| 第七节 对角矩阵和矩阵的约当标准型..... | 117 |
| 一、 n 阶方阵 A 变换成对角矩阵..... | 117 |
| 二、约当标准型..... | 121 |
| 第八节 状态转移矩阵的计算..... | 125 |
| 一、凯利-哈密顿定理 | 126 |
| 二、用凯利-哈密顿定理计算矩阵函数 | 127 |
| 三、 e^{At} 的近似计算 | 129 |
| 四、用 z 变换和拉氏变换计算状态转移矩阵..... | 129 |
| 第九节 系统的实现..... | 131 |

| | |
|------------------------------|------------|
| 一、由差分方程得到状态变量表达式 | 132 |
| 二、状态图和传递函数的实现 | 138 |
| 三、抽样数据系统的状态图和状态变量表达式 | 145 |
| 第十节 非均匀抽样数据系统的状态变量分析法 | 148 |
| 一、用状态变量分析法求抽样数据系统抽样时刻之间的响应 | 149 |
| 二、非均匀抽样数据系统抽样时刻的响应 | 150 |
| 附录 | 153 |
| 4.1 矩阵的定义 | 153 |
| 4.2 矩阵的基本运算 | 155 |
| 4.3 矩阵的秩 | 158 |
| 4.4 分块矩阵 | 159 |
| 4.5 矩阵行列式的运算 | 161 |
| 4.6 向量组的线性相关性 | 163 |
| 4.7 线性方程组 | 165 |
| 4.8 二次型与正定型 | 168 |
| 4.9 矩阵的微分和积分 | 171 |
| 习题 | 175 |
| 第五章 能控性和能观测性 | 181 |
| 第一节 概述 | 181 |
| 第二节 能控性 | 182 |
| 一、时不变系统 | 182 |
| 二、时变系统 | 187 |
| 第三节 能观测性 | 188 |
| 一、时不变系统 | 188 |
| 二、时变系统 | 191 |
| 第四节 系统状态完全能控和完全能观测的其它形式 | 192 |
| 一、同特征值相联系的状态完全能控条件 | 192 |
| 二、同特征值相联系的完全能观测条件 | 200 |
| 第五节 状态完全能控性和完全能观测性同传递函数的关系 | 204 |
| 第六节 对偶原理 | 207 |
| 第七节 系统完全能控和完全能观测的标准型 | 208 |
| 一、完全能控标准型 | 208 |
| 二、完全能观测标准型 | 210 |
| 附录 | 212 |
| 5.1 非奇异的相似变换不改变系统的状态完全能控性 | 212 |
| 5.2 非奇异的相似变换不改变系统的完全能观测性 | 213 |
| 习题 | 213 |
| 第六章 稳定性 | 217 |
| 第一节 概述 | 217 |
| 第二节 稳定性的定义 | 217 |

| | |
|---|-----|
| 第三节 数字系统的稳定判据 | 221 |
| 一、鞠利稳定性判据 | 221 |
| 二、双线性变换 | 224 |
| 三、劳斯-赫尔维茨判据 | 226 |
| 第四节 李亚普诺夫第二方法 | 228 |
| 一、李亚普诺夫稳定判据和不稳定判据 | 228 |
| 二、用李亚普诺夫第二方法设计最优状态反馈 | 232 |
| 附录 证明 A 的多项式 $f(A)$ 的特征值是 $f(\lambda_i)$ (λ_i 是 A 的特征值) | 234 |
| 习题 | 235 |
| 第七章 时域分析 | 237 |
| 第一节 概述 | 237 |
| 第二节 抽样周期对抽样数据系统输出波形和系统稳定性的影响 | 237 |
| 一、对输出波形的影响 | 237 |
| 二、对系统稳定性的影响 | 243 |
| 第三节 抽样数据系统瞬态响应同零极点位置的关系 | 244 |
| 一、 s 平面的极点同 z 平面的极点的映射关系 | 244 |
| 二、抽样数据系统的阶跃响应 | 245 |
| 第四节 z 平面上的根轨迹 | 257 |
| 第五节 稳态误差 | 261 |
| 一、定义 | 261 |
| 二、抽样数据系统的稳态误差 | 262 |
| 三、抽样周期对稳态误差的影响 | 266 |
| 第六节 数字仿真 | 267 |
| 一、数值积分 | 267 |
| 二、 z 型逼近法 | 272 |
| 附录 二阶连续系统的单位阶跃响应 | 275 |
| 习题 | 277 |
| 第八章 频域分析 | 282 |
| 第一节 概述 | 282 |
| 第二节 奈魁斯特判据 | 282 |
| 一、 z 变换法 | 284 |
| 二、无穷级数法 | 286 |
| 三、双线性变换 | 288 |
| 第三节 伯德图 | 289 |
| 第四节 增益裕量和相位裕量 | 293 |
| 一、增益裕量 | 294 |
| 二、相位裕量 | 294 |
| 第五节 闭环频率特性 | 295 |
| 一、对数幅相图和尼柯尔斯图线 | 295 |
| 二、频带宽度 | 297 |

| | |
|----------------------------|-----|
| 附录 | 299 |
| 8.1 奈魁斯特判据 | 299 |
| 8.2 计算 $GH(z)$ 的程序 | 305 |
| 8.3 尼柯尔斯图线 | 307 |
| 习题 | 309 |
| 第九章 数字控制系统的应用 | 311 |
| 第一节 概述 | 311 |
| 第二节 数字控制器的实现 | 313 |
| 一、用无源网络实现数字控制器 | 313 |
| 二、雷格悉尼(Ragazzini)直接设计法 | 318 |
| 三、用数字计算机实现数字控制器 | 321 |
| 第三节 相位超前和相位滞后数字控制器 | 323 |
| 一、频率特性 | 323 |
| 二、相位超前校正器 | 326 |
| 三、相位滞后校正器 | 330 |
| 第四节 比例积分微分(PID)数字控制器 | 331 |
| 一、数字 PID 控制器的 z 传递函数 | 332 |
| 二、用双线性变换法设计数字 PID 控制器 | 333 |
| 三、数字 PID 控制器的直接设计 | 337 |
| 第五节 根轨迹设计法 | 338 |
| 一、相位滞后控制器 | 339 |
| 二、相位超前控制器 | 341 |
| 第六节 最快响应数字控制系统的应用 | 342 |
| 一、闭环 z 传递函数的基本形式 | 343 |
| 二、对 $F(z)$ 和 $M(z)$ 的约束 | 346 |
| 第七节 无振荡数字控制系统的应用 | 350 |
| 一、单位阶跃输入 | 351 |
| 二、单位斜坡输入 | 354 |
| 三、用加权因子修改闭环 z 传递函数 | 355 |
| 四、特征方程的根对参数变化的敏感度 | 357 |
| 附录 用双线性变换法设计 PID 控制器的另一组公式 | 358 |
| 习题 | 359 |
| 第十章 极点配置和状态观测器 | 362 |
| 第一节 概述 | 362 |
| 第二节 求反馈矩阵 G 的其它方法 | 362 |
| 一、卡尔曼法 | 362 |
| 二、阿克曼公式 | 365 |
| 第三节 初始状态的转移 | 368 |
| 第四节 多输入极点配置 | 369 |
| 一、求反馈矩阵 G | 369 |

| | |
|---|------------|
| 二、不完全状态反馈..... | 371 |
| 第五节 输出反馈..... | 372 |
| 一、单输入..... | 372 |
| 二、多输入..... | 373 |
| 第六节 有输入的系统..... | 375 |
| 一、控制变量是状态变量和输入的线性组合..... | 375 |
| 二、具有状态反馈和动态输出反馈的系统..... | 377 |
| 第七节 状态观测器..... | 381 |
| 一、开环观测器..... | 381 |
| 二、闭环观测器..... | 381 |
| 三、用于观测器的阿克曼公式..... | 382 |
| 四、被控对象和观测器组成组合系统..... | 385 |
| 五、将状态观测器和反馈矩阵 \mathbf{G} 看成是一个控制器 | 388 |
| 第八节 降阶状态观测器..... | 390 |
| 第九节 输入对组合系统动态特性的影响..... | 392 |
| 一、求控制器的零点..... | 392 |
| 二、选择 \mathbf{M} 和 \mathbf{N} | 393 |
| 习题..... | 398 |
| 第十一章 线性二次型最优控制..... | 401 |
| 第一节 概述..... | 401 |
| 第二节 最优性原理与动态规划..... | 403 |
| 一、最优化原理..... | 403 |
| 二、线性二次型最优控制问题的一般解..... | 405 |
| 三、黎卡蒂增益矩阵的性质..... | 409 |
| 第三节 庞特里雅金最小值原理..... | 410 |
| 第四节 时间最优控制..... | 414 |
| 第五节 线性最优调节器..... | 417 |
| 一、有限终止时间线性最优调节器..... | 418 |
| 二、无限终止时间线性最优调节器..... | 419 |
| 习题..... | 425 |
| 第十二章 微处理器在数字控制系统中的应用..... | 428 |
| 第一节 概述..... | 428 |
| 第二节 截尾误差和舍入误差..... | 428 |
| 一、截尾误差..... | 428 |
| 二、舍入误差..... | 430 |
| 第三节 舍入效应对线性系统的影响..... | 431 |
| 一、最大输出误差..... | 431 |
| 二、稳态误差..... | 432 |
| 三、参数的舍入效应..... | 434 |
| 第四节 极限环..... | 436 |

| | |
|-----------------------------------|-----|
| 第五节 用微处理器的控制系统的时延..... | 437 |
| 第六节 举例..... | 438 |
| 一、用微处理器作数字控制器的直流电动机转速控制系统..... | 438 |
| 二、用微处理器作数字控制器的天线抛物面反射器方位控制系统..... | 441 |
| 附录 十六位 8086 微处理器简介 | 442 |
| 习题..... | 447 |
| 参考资料..... | 448 |

第一章 引 论

第一节 概 述

众所周知，数字电子计算机已广泛用来控制各种系统，例如，车辆、船舶、通信系统、飞机的自动驾驶仪以及各种机电伺服机构等。近年来，微处理器和微型计算机的迅速发展，为数字控制系统开辟了新的研究领域。

图 1.1 是数字控制系统的方框图，其中数字电子计算机相当一个数字控制器，它的数字输出保证系统达到给定的动态特性。图中 $r(t)$ 是连续参考输入， $c(t)$ 是连续受控信号（或称输出信号），它通过传感器产生负反馈信号 $b(t)$ 。起控制作用的连续误差信号 $e(t)$ 是 $r(t)$ 和 $b(t)$ 之差。A/D 变换器将 $e(t)$ 转换成计算机的数字输入信号 $m(kT)$ 。这里，假设这些数字信号具有固定的抽样周期 T （实际上，数字控制系统的抽样周期有时是可变的，有时不同反馈环路具有不同的周期^[1]）。计算机的逻辑系统包含时钟脉冲部分，它每隔 T 秒产生一个脉冲。A/D 变换器按照时钟脉冲的周期不断地向计算机输送数字序列 $m(kT)$ ，它经过计算机的处理，得到输出数字序列 $u(kT)$ ，再经 D/A 变换器，产生连续控制信号 $u(t)$ ，对被控对象实现控制。

图 1.1 是一个闭环控制系统，它依靠指示误差信号 $e(t)$ 起控制作用，例如实际输出值 $c(t)$ 若大于理想输出值，则传感器产生的反馈信号 $b(t)$ 加大，这使 $e(t)$ 减小，从而使 $c(t)$ 的值下降。反之亦然。

数字控制的一个主要目的是使被控对象具有满意的动态响应。即使被控对象有干扰 $w(t)$ （见图 1.1），传感器有误差 $v(t)$ ，系统的输出 $c(t)$ 也应能跟踪 $r(t)$ 的变化。具有良好调节作用的系统，对外界干扰有好的抑制能力，对被控对象参数的变化具有低的敏感度。

图 1.1 的数字控制系统既有数字信号，又有连续信号，也称抽样数据系统。

由于有反馈的控制系统比无反馈的控制系统具有更大的优越性，本书主要讨论有反馈的数字控制。

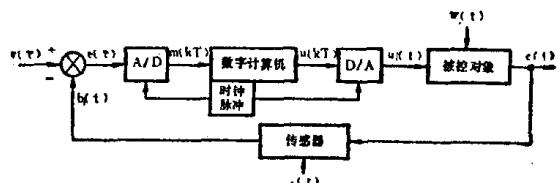


图 1.1 数字控制系统的方框图

第二节 数字控制系统的优点

同连续控制系统相比，数字控制系统的主要优点是：

(1) 作为控制系统中数字控制器的数字电子计算机可以多路复用，这可用图 1.2 说明，工作时，输入通道 1 的信号首先通过多路转换器 1，经数字电子计算机处理，在多路转换器 2 的输出通道 1 得到输出信号。在一定的时间间隔后，多路转换器 1 将输入通道 2 接通，按同

[1] 见第三章第九节和第四章第十节。

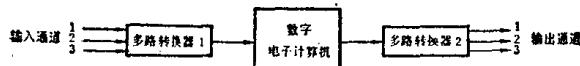


图 1.2 多路复用数字控制器系统

样步骤，在输出通道 2 产生经计算机处理过的输出信号。再经过同样的时间间隔，输出通道 3 上有输出信号。这样一来，每一通道都好像单独使用一个数字电子计算机。数字电子计算机的速度愈快，能同时处理的通道愈多，设备的利用率也愈高。

(2) 使用灵活：改变数字控制器的程序可以改变它的特性，因此无需对数字系统的硬件设备作任何改变，就可以改变控制系统的性能。

(3) 可靠性高：数字系统采用的电子元件、传感器和编码器等比较可靠，体积较小，也较少受环境条件变化的影响。

由于上述优点，控制系统已日益广泛地采用数字控制器以及相应的数模转换设备。

数字系统的缺点是，它的速度还不很高，例如 A/D 变换器、抽样器和逻辑硬件只能工作在几十兆赫以下，不能处理频率很高的信号。

第二章 信号的变换

第一节 概述

前已述及，在数字控制系统中广泛采用数字电子计算机作为数字控制器，而数字电子计算机的信号是用数码表示的。用 A/D 变换器，可以将模拟信号转换成数字电子计算机需要的数字信号。为了控制被控对象，又需利用 D/A 变换器将数字信号转换成模拟信号。这种模-数和数-模信号的转换过程称为信号的变换。

抽样器是将模拟信号转换成抽样信号的一种设备，它的作用是每隔一定的时间间隔，对模拟信号抽样。为了使抽样信号能够包含模拟信号的全部信息，抽样的速率必须足够高。分析证明，当抽样率大于或等于模拟信号最高频率的两倍时，就可达到上述目的。

在信号变换中，广泛采用保持器，它有零阶、一阶和高阶之分。常用的零阶保持器的作用是，当输入是脉冲序列时，保持器的输出在抽样周期 T 内保持不变。保持器可用来将数字信号恢复成模拟信号，也可以接在抽样器后面保证 A/D 变换器的正常工作。

第二节 数字信号和量化

一、数字信号

数字电子计算机的数字信号通常用二进位计数制(简称二进制)来表示。在二进制中，只有两个数字“0”和“1”，并且由低位向高位的进位是“逢二进一”。一系列二进制数 0 和 1 组成“字”，字的长度(即包含的位数)，称“字长”^[1]。字长因不同计算机而异，大型计算机的字长可达 32 位，而微型计算机的字长通常是 4、8 或 16 位。通常将 8 位称作一个字节。对于 8 位微型计算机，字长是一个字节。

数字电子计算机的精度取决于字长，字长加大，精度提高，机器的复杂程度也随之增加。

一个无符号 n 位二进制数的数值为

$$N = a_{n-1} 2^{n-1} + \cdots + a_2 2^2 + a_1 2^1 + a_0 2^0 \quad (2.1)$$

式中数码 a_i ($i=0, 1, \dots, n-1$) 等于 0 或 1，最左位称作最高位 (MSB)，最右位称作最低位 (LSB)。同一数码在不同的数位所代表的数值是不同的。每一数位代表的常数称数位权值，简称位权。例如，最低位的位权是 2^0 ，最高位的位权是 2^{n-1} 。当 $a_i=1$ 时，相应的数值等于 1 乘以位权 2^i ； $a_i=0$ 时，数值等于零。例如

$$N = a_3 2^3 + a_2 2^2 + a_1 2^1 + a_0 2^0$$

当 a_0 、 a_1 、 a_2 和 a_3 是不同的 0 和 1 的组合时，数 N 可表示 16 个(2^4 个)不同的状态或数。

[1] 计算机的字长指的是机器一次并行处理的二进制位数。

在计算机中，数的表示方法有两种：定点表示法和浮点表示法。定点表示法指的是，数的小数点^[1]位置是固定的；而浮点表示法的小数点位置是可以变动的。

图2.1表示带小数部分的8位数的定点表示。如二进制数是11010.001，它相当于十进制数。

$$\begin{aligned} N &= 1 \times 2^4 + 1 \times 2^3 + 0 \times 2^2 + 1 \times 2^1 + 0 \times 2^0 + 0 \times 2^{-1} + 0 \times 2^{-2} + 1 \times 2^{-3} \\ &= 16 + 8 + 2 + \frac{1}{8} = 26.125 \end{aligned}$$

一般，可将 n 位小数表示为

$$N = a_{-1}2^{-1} + a_{-2}2^{-2} + \cdots + a_{-n}2^{-n} \quad (2.2)$$

式中 $a_i (i = -1, -2, \dots, -n)$ 代表 0 或 1。

上面讨论了无符号数的二进制表示法。为了表示负数，可用最高位代表符号位^[2]。当符号位是 0 时表示正数，是 1 时表示负数。设有一个三位字，它的最高位是符号位，该字能表示的最大整数是 $2^2 - 1 = 3$ ，而最小整数是 $-(2^2 - 1) = -3$ 。一般说，具有符号位的 n 位字能表达的整数，位于 $(2^{n-1} - 1)$ 和 $-(2^{n-1} - 1)$ 之间。

对于具有符号位的 n 位字，如果小数点后有 m 位，则它能表示的数位于 $(2^{n-1} - 1)2^{-m}$ 和 $-(2^{n-1} - 1)2^{-m}$ 之间。

定点表示法的优点是计算机设备比较简单，缺点是能表达的数的范围较小，当运算的结果超出字长所能表达的数时，将产生“溢出”，计算机不能正常工作。

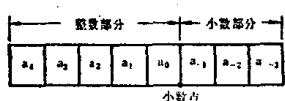


图 2.1 具有小数部份的 8 位数的定点表示

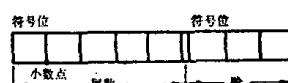


图 2.2 8 位数的浮点表示法

浮点表示法又称科学标志法，它表示数的方法示如图 2.2。图中 8 位数的第一部分是“尾数”，第二部分是“阶”。同样一个数，可以用不同的尾数和阶来表示，这相应于改变小数点的位置。例如，十进制数 6，可以写成 0.6×10^1 （尾数是 0.6，阶是 1），也可写成 600×10^{-2} （尾数是 600，阶是 -2）。

在计算机中，二进制浮点数可表示为

$$N = M \times 2^E \quad (2.3)$$

式中 M 是尾数， E 是数 N 的指数的阶。通常，尾数采用规格化的形式，也即，尾数代表的十进制数的值必须满足

$$0.5 \leq M < 1$$

考虑到尾数和阶可以是正或负的，将尾数和阶的最高位作为符号位，如图 2.2 所示。对于规格化形式的尾数，紧接在符号位后面的第一位是 1，小数点位于符号位和 1 之间（图 2.2）。对于字长较短的微型计算机，可以用两个连接的字构成浮点数，如图 2.3 所示。

式(2.3)中的阶 E 代表小数点移动的位数，当 $E > 0$ 时，小数点右移，当 $E < 0$ 时，小数点左移。例如，图 2.4 表示十进制数 7 的 8 位浮点二进制字，其中尾数占 5 位，符号位是 0，小数部分占 4 位，表示成 1110；阶占 3 位，符号位是 0，两位二进制整数 11 代表十进制 3。因此，它表示的数 N 是

[1] 计算机中数的小数点，并不用任何记号表示出来，而是采取人为约定的办法。

[2] 有符号位的数称有符号数。

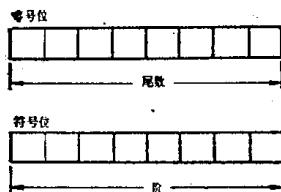


图 2.3 用两个字表示浮点数

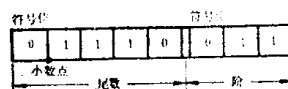


图 2.4 用 8 位浮点二进制字表示十进制数 7

$$N = \cdot 1110 \times 2^{11} (\text{二进制})$$

$$= \left(\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + 0 \right) 2^3 (\text{十进制}) = 7$$

上式说明，将·1110的小数点右移三位，可得十进制 7。

下面求出浮点表示法能表示的最大数和最小数。设有 n 位字，尾数占 m 位（包括符号位），阶占 e 位（包括符号位），能表达的最大数示如图 2.5 a，其中，除去两个符号位都是 0 以外，其它各位都是 1。经过简单的计算，可以求出最大数 N_{\max}

$$N_{\max} = (1 - 2^{-m+1}) 2^{(2^{e-1}-1)} \quad (2.4)$$

例如，对于图 2.4 所示的 8 位浮点字， $m=5$, $e=3$

$$N_{\max} = (1 - 2^{-4}) 2^{(2^2-1)} = \left(1 - \frac{1}{16}\right) 2^3 = 7.5$$

如果尾数和阶各占 4 位，即 $m=4$, $e=4$ ，有

$$N_{\max} = (1 - 2^{-3}) 2^{(2^3-1)} = \left(1 - \frac{1}{8}\right) 2^7 = 112$$

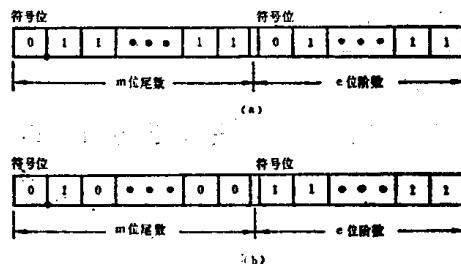


图 2.5 浮点数的最大和最小幅值

表 2.1 八位二进制表示数的范围

| 十进制数 | 补码 |
|------|----------|
| 127 | 01111111 |
| ⋮ | ⋮ |
| 1 | 00000001 |
| 0 | 00000000 |
| -1 | 11111111 |
| ⋮ | ⋮ |
| -126 | 10000010 |
| -127 | 10000001 |
| -128 | 10000000 |

图 2.5 (b) 表示具有 n 位字的最小正数 N_{\min} ，这时，尾数符号位后第一位是 1，其它均为 0；阶的符号位为 1，其它各位也是 1。 N_{\min} 是

$$N_{\min} = 0.5 \times 2^{-(2^{e-1}-1)} \quad (2.5)$$

浮点法的优点是表示数的范围较大，并且工作比较方便。缺点是机器比较复杂，运算速度比定点法慢。

在计算机中，用原码、反码和补码表示有符号数。原码是用符号位表示数的符号，用数的绝对值表示数值部分。例如 $+9/16$ 的原码为 0.1001， $-9/16$ 的原码为 1.1001。正数的反码同原码一样；负数的反码是将其原码的各位（除去符号位）求反。例如， $-9/16$ 的反码是

$$\begin{array}{ccc} 1.1001 & \xrightarrow{\text{求反}} & 1.0110 \\ \text{原码} & & \text{反码} \end{array}$$

正数的补码也同原码一样；负数的补码是反码的最低位加 1。例如， $-9/16$ 的补码是

$$\begin{array}{ccccc} 1.1001 & \xrightarrow{\text{求反}} & 1.0110 & \xrightarrow{\text{最低位加1}} & 1.0111 \\ \text{原码} & & \text{反码} & & \text{补码} \end{array}$$