排灌机械工程学报
Journal of Drainage and Irrigation Machinery Engineering
2016 年 2 月 Feb. 2016 | 第 34 卷 第 2 期 Vol. 34 No. 2
doi:10.3969/j.issn.1674-8530.15.0108

投影寻踪法在泵站建筑物安全评价中的应用


曹睿哲

曹睿哲[1,2]，田浪[1,2]，郭蓉[1,2]，肖俊龙[1,2]
（1.河海大学南方地区高效灌排与农业水土环境教育部重点实验室，江苏 南京 210098；
2.河海大学水利水电学院，江苏 南京 210098）

摘要：为了评价在役泵站工程建筑物安全状况，建立了泵站工程建筑物安全评价指标体系.对安全评价指标重要性进行分析，着重分析了泵站工程建筑物安全评价客观赋权方法之一的投影寻踪方法.根据泵站工程建筑物安全评价指标的实测值求计算数据，通过投影寻踪法模型计算并联系基于遗传算法的投影寻踪计算原理分析计算结果.利用江苏省某泵站作为实例，对影响泵站建筑物安全状态各指标层指标重要性进行分析，分别对混凝土结构安全状态和站身稳定安全状态 2 个指标层进行了研究.利用投影寻踪计算模型分析各指标层下各评价指标重要性，应用 Matlab 自带的遗传算法工具箱进行具体计算.计算结果表明，此泵站混凝土结构安全状态指标层中混凝土实测强度最为重要，而泵站站身稳定安全指标层中站身抗滑稳定这一指标权重最大.

关键词：泵站；客观赋权；遗传算法；投影寻踪法；权重计算
中图分类号：TV698.1　**文献标志码**：A　**文章编号**：1674-8530(2016)02-0133-05

曹睿哲，田浪，郭蓉，等. 投影寻踪法在泵站建筑物安全评价中的应用[J]. 排灌机械工程学报，2016，34(2)：133-137.

CAO Ruizhe，TIAN Lang，GUO Rong，et al. Application of projection pursuit method in pumping station safety evaluation[J]. Journal of drainage and irrigation machinery engineering(JDIME)，2016，34(2)：133-137.(in Chinese)

Application of projection pursuit method in pumping station safety evaluation

CAO Ruizhe[1,2]，*TIAN Lang*[1,2]，*GUO Rong*[1,2]，*XIAO Junlong*[1,2]
(1. Key Laboratory of Efficient Irrigation-Drainage and Agriculture Soil-Water Environment in Southern China，Hohai University，Nanjing，Jiangsu 210098，China；2. College of Water Conservancy and Hydropower，Hohai University，Nanjing，Jiangsu 210098，China)

Abstract：In order to evaluate the in-service pumping station building security status，a pumping station building safety evaluation system is built. The importance of safety evaluation index is analyzed，and the projection pursuit method which is an objective weighting method of pumping station building safety evaluation is specially analyzed. Data are calculated according to the measured value of the pumping station building safety evaluation index. It is calculated through the projection pursuit method model，and the calculation results are analyzed，combined with the principle of projection pursuit method based on genetic algorithm. A pumping station in Jiangsu Province is used as an example. The index importance of each index layer of the pumping station building security status is studied. Two indicator layers：concrete structure safety status and station body stability safety status indicators layer are studied. The projection pursuit model is used to analyze the importance of each evaluation index under

收稿日期：2015-05-13；**网络出版时间**：2016-01-25
网络出版地址：http://www.cnki.net/kcms/detail/32.1814.TH.20160125.1057.016.html
基金项目：国家自然科学基金重点资助项目(41323001)
作者简介：曹睿哲(1991—)，男，江苏扬州人，硕士研究生(347196883@qq.com)，主要从事水工建筑物安全评价研究.
田浪(1993—)，男，湖北仙桃人，硕士研究生(1542450186@qq.com)，主要从事水利工程建设与管理研究.

each indicator layer. Matlab genetic algorithm toolbox is applied to compute each evaluation index weight. The results show that the measured strength of concrete is most important in the pumping station concrete structure safety status indicator layer, and the index weight of stability against sliding in the pumping station stand body stability safety index layer is the largest.

Key words: pumping station; objective weighting method; genetic algorithm; projection pursuit method; weighting calculation

泵站工程是水利工程的重要组成部分,在国民经济的区域性灌溉、排涝、调水、供水方面有着重要的作用,但是目前中国泵站建筑物由于设计标准低、年久失修、管理不到位等原因,普遍存在建筑物及设备老化等问题[1-2],因此,对泵站工程进行安全综合评价具有重要的现实意义,而在安全评价环节中,评价指标的权重计算又是一项重要的工作.

目前确定权重的分析方法主要分为主观赋权法和客观赋权法2类[3-4]. 主观赋权法主要是依据专家的经验,根据对权重指标的重视程度主观上确定权重指标的赋权方法,如层次分析法、专家咨询法、二项系数法等. 主观赋权法的权重受决策者的主观意愿的影响很大,决策者往往依据自己的经验来确定权重值,解释性较强,但客观性较差. 客观赋权法则是通过处理各指标的原始信息,采用数学的方法,计算出指标的权重. 客观赋权法确定的权重在大多数情况下精度较高. 文中研究客观权重的赋权方法之一的投影寻踪法,并通过对工程实例的分析,探讨投影寻踪法的合理性.

1 计算原理与方法

1.1 计算原理

遗传算法(genetic algorithm,GA)是一种模拟生物界遗传和进化过程而建立的搜索算法,其基本思路是从一组随机产生的初始"种群"解开始,进行搜索,种群中的每一个问题的解,称为"染色体";遗传算法是通过染色体的"适应值"评价染色体的好坏,染色体的适应值越大,被选择的概率越高,相反,染色体的适应值越小,被选择的可能性越小,被选择后的染色体随即进入下一代;下一代中的染色体又通过交叉和变异等遗传操作,继续产生新的染色体,即"后代";最后经过若干代,算法收敛于最好的染色体,最终的染色体就是问题的最优解或近似解.

GA利用简单的编码技术和遗传操作算子模拟复杂的优化过程,可以解决非常困难的优化问题. 该算法只要求优化问题为可计算的,突破了问题结构是否为线性、连续、可微、单峰、噪声等各种因素限制,同时,也不受优化准则函数的形式、优化变量数目的束缚,对搜索空间也没有特殊要求(如要求为凸性、连通性),它可直接在适应度函数值的引导下,在搜索空间进行自适应的全局搜索,运行过程简单,但计算结果丰富,是目前优秀、通用性强的一种优化方法,已在许多实际优化问题中得到广泛应用和推广[5].

遗传算法优化方法是模拟智能过程的一类新颖优化方法,与常规优化方法相比,遗传算法优化的结果是一组好的解而不是单个解,这为该算法的使用者提供了可供选择的机会,因此,遗传算法优化方法特别适合处理复杂的系统评价问题[6].

工程上常用的一种GA是标准遗传算法(SGA)[7],模型的参数优化问题为

$$\min f = \sum_{i=1}^{m} \| F(C, \boldsymbol{X}_i) - \boldsymbol{Y}_i \|^q$$
$$\text{s.t. } a_j \leqslant c_j \leqslant b_j, (j = 1, 2, \cdots, p), \quad (1)$$

式中:f为优化准则函数;$C = \{c_j\}$为模型p个待优化参数(优化变量);$[a_j, b_j]$为c_j的初始变化区间(搜索区间);$\boldsymbol{X}$为模型N维输入向量;$\boldsymbol{Y}$为模型M维输入向量;F为一般非线性模型,即$F: R^N \rightarrow R^M$;$\| \ \|$为取范数;q为实常数,q取1时为最小一乘准则,q取2时为最小二乘准则等,可视实际的建模要求而定.

KRUSKAL首次提出投影寻踪法[8],把数据投影到低维子空间,通过极大化默写指标,以发现数据的聚类结构. 最初投影寻踪方法仅作为一种数值计算的方法出现,并且仅限于做判别分析和聚类分析. 随着投影寻踪思想的推广,很多学者逐步提出了投影寻踪回归、基于数据的矩投影指标、投影寻踪密度估计等方法. 众多国内外学者的研究成果表明,作为一类新兴的统计方法,投影寻踪方法具有很强的生命力及广阔的应用前景.

投影寻踪法是一类新兴的多元数据分析的数据方法,它是用来处理和分析高维数据,尤其是来自非正态总体分布的高维数据的一种探索性分析的有效方法,其基本思想是把高维数据通过某种组

合，投影到低维（1－3 维）子空间上，通过极大（小）化某个投影指标，寻找出能反映高维数据结构或特征的投影，在低维空间上对数据结构进行分析，以达到研究和分析高维数据的目的.

1.2　计算步骤

基于遗传算法优化的投影寻踪法[9－11]计算指标的权重的具体步骤如下：

1）样本评价指标集的归一化处理.

设评价等级表中各指标值的样本集为$\{x^*(i,j)\mid i=1,2,\cdots,n,j=1,2,\cdots,p\}$，其中$x^*(i,j)$为第$i$个样本第$j$个指标值，$n,p$分别为样本的个数（样本容量）和指标的数目. 为消除各指标量纲和统一各指标变化范围，可先将各指标值归一化处理，则形成的决策矩阵$\boldsymbol{x}$为

$$x(i,j)=\begin{bmatrix} & D_1 & D_2 & \cdots & D_j \\ M_1 & x_{11} & x_{12} & \cdots & x_{1j} \\ M_2 & x_{21} & x_{22} & \cdots & x_{2j} \\ \vdots & \vdots & \vdots & & \vdots \\ M_i & x_{i1} & x_{i2} & \cdots & x_{ij} \end{bmatrix}, \tag{2}$$

其中，$x(i,j)$为指标值归一化的序列.

2）构造指标函数$Q(a)$.

投影寻踪法就是把p维数据$\{x^*(i,j)\mid j=1,2,\cdots,p\}$综合成以$a=(a(1),a(2),a(3),\cdots,a(p))$为投影方向的一维投影值$z(i)$，即

$$z(i)=\sum_{j=1}^{p}a(j)x(i,j),i=1,2,\cdots,n, \tag{3}$$

然后根据$\{z(i)\mid i=1,2,\cdots,n\}$的一维散布图进行分类，其中，$a$为单位长度向量. 综合投影指标值时，要求投影值$z(i)$的散布特征应为局部投影尽可能密集，最好凝聚成若干个点团，而在整体上投影点团之间尽可能散开. 因此，投影指标函数表达为

$$Q(a)=S_zD_z, \tag{4}$$

其中，S_z为投影值$z(i)$的标准差；D_z为投影值$z(i)$的局部密度，即

$$S_z=\sqrt{\frac{\sum_{i=1}^{n}(z(i)-E(z))^2}{n-1}}, \tag{5}$$

$$D_z=\sum_{i=1}^{n}\sum_{j=1}^{n}(R-r(i,j))\mu(R-r(i,j)), \tag{6}$$

式中：$E(z)$为序列$\{z(i)\mid i=1,2,\cdots,n\}$的平均值；$R$为局部密度的窗口半径，它的选取既要使包含在窗口内的投影点的平均个数不太少，避免滑动平均偏差太大，又不能使它随着n的增加而太高，R可以根据试验确定，其取值范围为$r_{\max}+\frac{p}{2}\leqslant R\leqslant 2p$，一般可取0.1倍的$S_z$. $r(i,j)$表示样本之间的距离，$r(i,j)=|z(i)-z(j)|$，$u(t)$为一单位阶跃函数，当$t\geqslant 0$时，其函数值为1；当$t<0$时，其函数值为0.

3）优化投影指标函数.

当各指标值的样本集合给定时，投影指标函数$Q(a)$只随着投影方向a的变化而变化，不同的投影方向反映不同的数据结构特征，最佳投影方向就是最大可能暴露高维数据某类特征结构的投影方向. 因此，可以通过求解投影指标函数最大化问题来估计最佳投影方向，即

$$\max Q(a)=S_zD_z, \tag{7}$$

$$\text{s.t.}\ \sum_{j=1}^{n}a^2(j)=1, \tag{8}$$

这是一个以$\{a(j)\mid j=1,2,\cdots,p\}$为优化变量的复杂非线性优化问题.

4）求解优化结果.

通过Matlab自带遗传算法工具箱[12－13]，根据实际问题编写待优化函数及约束条件的M文件，并通过设置各遗传参数等，计算基于遗传算法优化的最佳投影方向，从而得出各指标的权重.

2　实例分析

2.1　工程概况

江苏省某泵站承担废黄河流域、黄墩湖滞洪区的部分灌溉和排涝任务，该站于1987年12月建成，安装900HL－10型立式混流泵配用JSL－15－10型电动机5台套，其额定功率和电压分别为260 kW和6 kV，总装机容量为4 570 kW，抽水能力为29 m^3/s.

2.2　站身建筑物安全评价指标体系构建

泵站建筑物的安全状态主要由工程现场安全检测、工程复核计算2大部分组成，泵站建筑物现场安全检测的内容主要包括混凝土结构强度、混凝土外观、裂缝、钢筋锈蚀、混凝土碳化深度和保护层厚度等. 泵站的安全复核内容包括消能防冲复核、抗渗稳定复核、整体稳定复核、地基应力复核、结构构件的承载力复核等.

基于评价指标构建的层次性、科学性、简洁性、相对独立性及可操作性原则，现对该站建立建筑物安全状态评价体系，并对泵站混凝土构件安全和站身稳定用投影寻踪法计算各评价指标权重.

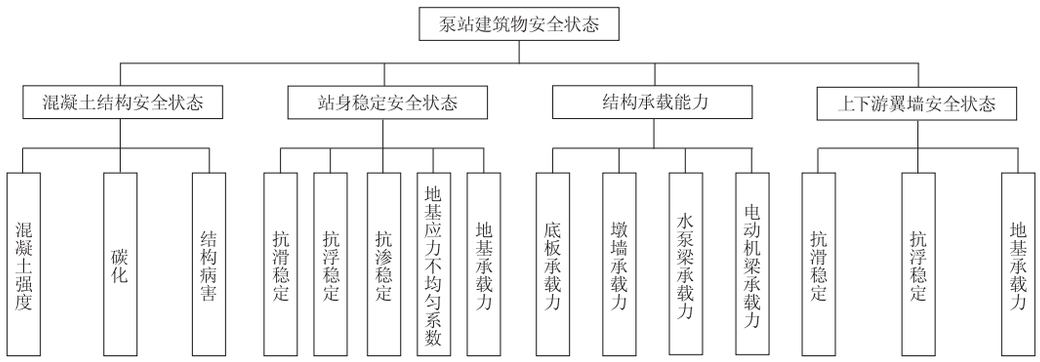


图1 某泵站泵房建筑物安全评价体系

Fig. 1 Pump room building safety evaluation system of a pumping station

2.3 定量评价指标的规格化

在泵站工程安全评价体系中,各定量指标在评价之前需进行量纲一化和规格化处理,文中泵站工程底层评价指标规格化参考已知的分级评价标准[5,14-15].各指标值及指标规格化之后的值见表1,2.

表1 某泵站建筑物混凝土结构安全评价指标及规格化

Tab. 1 Building concrete structure safety evaluation index and specification of a pumping station

构件名称	指标 a_1		指标 a_2		指标 a_3
	碳化系数	指标规格化	检测强度	指标规格化	结构病害
1#机组左墩	0.921	0.837	19.4	0.394	0.80
1#机组右墩	0.662	0.730	21.8	0.544	0.65
3#机组左墩	0.857	0.825	19.9	0.399	0.71
3#机组右墩	0.621	0.696	24.6	0.768	0.68
5#机组左墩	0.751	0.800	19.9	0.399	0.70
5#机组右墩	0.962	0.844	22.2	0.576	0.65
1#机组电机梁	0.021	0.408	28.2	0.823	0.45
3#机组电机梁	-0.063	0.387	26.1	0.808	0.35
5#机组电机梁	-0.696	0.261	22.6	0.608	0.50
1#机组出水管	0.050	0.420	35.1	0.858	0.85
3#机组出水管	0.766	0.804	35.1	0.858	0.80
5#机组出水管	0.822	0.818	32.7	0.847	0.80
1#机组泵房前墙	0.310	0.524	18.8	0.388	0.65
3#机组泵房前墙	0.840	0.822	28.1	0.822	0.70
5#机组泵房前墙	0.243	0.497	18.6	0.386	0.65
1#机组出水挡土墙	-0.259	0.348	21.0	0.480	0.55
3#机组出水挡土墙	0.559	0.647	22.1	0.568	0.45
5#机组出水挡土墙	0.648	0.719	22.2	0.576	0.60

表2 某泵站建筑物站身稳定安全评价指标及规格化

Tab. 2 Building stand body stability safety evaluation index and specification of a pumping station

工况	指标 b_1		指标 b_2	
	计算值	规格化值	计算值	规格化值
1	1.30	0.880	1.97	0.888
2	1.11	0.802	1.85	0.881
3	1.43	0.860	2.10	0.904
4	1.23	0.824	1.89	0.883
5	0.87	0.400	1.90	0.884

续表

工况	指标 b_3		指标 b_4		指标 b_5	
	计算值/m	规格化值	计算值	规格化值	计算值/kPa	规格化值
1	51.0	0.816	1.21	0.879	47.78	0.906
2	54.0	0.817	1.54	0.846	46.34	0.908
3	46.2	0.817	1.06	0.894	48.77	0.904
4	49.2	0.824	1.14	0.886	43.93	0.914
5	51.0	0.806	2.95	0.783	45.93	0.910

2.4 计算结果分析

根据基于遗传算法优化的投影寻踪法的计算步骤,分别计算A,B这2个二级指标下的各底层指标的权重,计算结果见表3.

表3 某泵站建筑物安全评价底层指标权重

Tab. 3 Building safety evaluation bottom index weight of a pumping station

二级指标	底层指标	权重值
A	指标 a_1	0.378 7
	指标 a_2	0.424 2
	指标 a_3	0.197 1
B	指标 b_1	0.903 7
	指标 b_2	0.021 3
	指标 b_3	0.020 7
	指标 b_4	0.034 0
	指标 b_5	0.020 3

由表3中计算结果可知,由原始数据分析在混凝土构件安全评价中,底层指标 a_1,a_2,a_3 的权重分别为0.378 7,0.424 2,0.197 1,指标 a_2 所占的权重最大,说明由实际检测数据中,混凝土结构强度指标对泵站建筑物的安全评价有重要的影响,混凝土碳化、结构病害的影响次之.在泵站站身稳定安全状态评价中,底层指标 b_1,b_2,b_3,b_4,b_5 权重分别为0.903 7,0.021 3,0.020 7,0.034 0,0.020 3,指标 b_1 所占的比重最大,为0.963 7,说明该项指标抗滑稳定的计算结果对泵站站身稳定安全状态有着重要的影响.因此,在实际泵站站身建设中,应十分重视该指标值.

3　结　论

1）基于遗传算法优化的投影寻踪法作为客观赋权方法的1种，其计算方法相对简单，灵敏度较高，其判断结果不依赖主观判断，有较强的数学理论依据，计算结果可为权重的确定提供可靠的参考.

2）由于在基于遗传算法优化的投影寻踪法的使用时需要依赖足够的样本数据，如没有足够的样本数据或遇到部分指标值数据间的变动很小而部分指标值中有较大突变情况时，计算指标权重会有一定的局限性.

3）针对江苏省某泵站混凝土结构安全状态和站身稳定安全状态进行分析，混凝土结构安全状态指标层中混凝土实测强度影响最大，而站身稳定安全状态指标层中站身抗滑稳定的权重最大，即站身抗滑稳定影响最大.

参考文献(References)

[1] 陈坚，李琪，许建中，等. 中国泵站工程现状及“十一五”期间泵站更新改造任务[J]. 水利水电科技进展，2008，28(2)：84－88.
CHEN Jian, LI Qi, XU Jianzhong, et al. Current status of pump stations in China and their renovation and rebuilding according to the 11th Five-Year Plan[J]. Advances in science and technology of water resources, 2008, 28(2): 84－88. (in Chinese)

[2] 刘竹溪，冯广志. 中国泵站工程[M]. 北京：水利电力出版社，1993.

[3] 郭亚军. 综合评价理论、方法及拓展[M]. 北京：科学出版社，2012.

[4] PULFORD J, ADAMS P, SHERIDAN J. Developing a clinical assessment model suited to use in an agency providing short-term substance use treatment: findings from a Delphi survey of expert opinion[J]. Administration & policy in mental health & mental health services research, 2009, 36(5): 322－330.

[5] 许文婷. 基于遗传算法的大型泵站工程安全综合评价模型研究[D]. 扬州：扬州大学，2012.

[6] 周玉良. 水安全问题中的智能非参数方法[D]. 合肥：合肥工业大学，2006.

[7] 金菊良，杨晓华，丁晶. 标准遗传算法的改进方案——加速遗传算法[J]. 系统工程理论与实践，2001，21(4)：8－13.
JIN Juliang, YANG Xiaohua, DING Jing. An improved simple genetic algorithm—accelerating genetic algorithm[J]. Systems engineering-theory & practice, 2001, 21(4): 8－13. (in Chinese)

[8] 田铮，林伟. 投影寻踪方法与应用[M]. 西安：西北工业大学出版社，2008.

[9] 金菊良，张欣莉，丁晶. 评估洪水灾情等级的投影寻踪模型[J]. 系统工程理论与实践，2002(2)：140－144.
JIN Juliang, ZHANG Xinli, DING Jing. Projection pursuit model for evaluating grade of flood disaster loss[J]. Systems engineering-theory & practice, 2002(2): 140－144. (in Chinese)

[10] 陈曜，丁晶，赵永红. 基于投影寻踪原理的四川省洪灾评估[J]. 水利学报，2010，41(2)：220－224.
CHEN Yao, DING Jing, ZHAO Yonghong. Assessment on flood disaster in Sichuan Province based on the principle of projection pursuit method[J]. Journal of hydraulic engineering, 2010, 41(2): 220－224. (in Chinese)

[11] 封志明，郑海霞，刘宝勤. 基于遗传投影寻踪模型的农业水资源利用效率综合评价[J]. 农业工程学报，2005，21(3)：66－70.
FENG Zhiming, ZHENG Haixia, LIU Baoqin. Comprehensive evaluation of agricultural water use efficiency based on genetic projection pursuit model [J]. Transactions of the CSAE, 2005, 21(3): 66－70. (in Chinese)

[12] 陈广洲，解华明，鲁祥友. Matlab 遗传算法工具箱在非线性优化中的应用[J]. 计算机技术与发展，2008，18(3)：246－249.
CHEN Guangzhou, XIE Huaming, LU Xiangyou. Nonlinear optimization based on genetic algorithm toolbox of Matlab[J]. Computer technology and development, 2008, 18(3): 246－249. (in Chinese)

[13] 何逢标. 综合评价方法 MATLAB 实现[M]. 北京：中国社会科学出版社，2010.

[14] 李贝. 大中型泵站工程的健康诊断方法研究[D]. 扬州：扬州大学，2010.

[15] 金秋. 混凝土坝健康诊断指标体系及指标重要性分析方法研究[D]. 南京：河海大学，2010.

（责任编辑　盛杰）