1997

Resolution of Local Inconsistency in Identification

Douglas Ray Anderson
Metro

Martin Zwick
Portland State University, zwick@pdx.edu

RESOLUTION OF LOCAL INCONSISTENCY IN IDENTIFICATION

Douglas Ray Anderson[#] and Martin Zwick[*]
Systems Science Ph.D. Program[&]
Portland State University
Portland OR 97027

Abstract

This paper reports an algorithm for the resolution of local inconsistency in information-theoretic identification. This problem was first pointed out by Klir as an important research area in reconstructability analysis. Local inconsistency commonly arises when an attempt is made to integrate multiple data sources, i.e., contingency tables, which have differing common margins. For example, if one has an AB table and a BC table, the B margins obtained from the two tables may disagree. If the disagreement can be assigned to sampling error, then one can arrive at a compromise B margin, adjust the original AB and BC tables to this new B margin, and then obtain the integrated ABC table by the conventional maximum uncertainty solution.

The problem becomes more complicated when the common margins themselves have common margins. The algorithm is an iterative procedure which handles this complexity by sequentially resolving increasingly higher dimensional inconsistencies. The algorithm is justified theoretically by maximum likelihood arguments. It opens up the possibility of many new applications in information theoretic modeling and forecasting. One such application, involving transportation studies in the Portland area, will be briefly discussed.

[#] 503-797-1788, andersond@metro.dst.or.us
[*] 503-725-4987, zwick@sysc.pdx.edu
[&] http://www.sysc.pdx.edu

Talk for IIGSS, 1997, San Marcos, Texas

1. IDENTIFICATION UNDER LOCAL INCONSISTENCY

PROBLEM: RESOLVING LOCAL INCONSISTENCY AMONG PROJECTIONS

2. CONFORM ALGORITHM

METHOD: OBTAIN COMPROMISE MARGINS WHICH MINIMIZE ERROR

3. APPLICATION

RESULTS: TRANSPORTATION MODELING & FORECASTING

1. IDENTIFICATION UNDER LOCAL INCONSISTENCY

PROBLEM: RESOLVING LOCAL INCONSISTENCY AMONG PROJECTIONS

- THE IDENTIFICATION PROBLEM

- EXAMPLE & ITS SOLUTION

- EQUATIONS

- IDENTIFICATION WITH LOCAL INCONSISTENCY

- EXAMPLE & ITS SOLUTION

- EQUATIONS

- GLOBAL INCONSISTENCY

2. CONFORM ALGORITHM

METHOD: OBTAIN COMPROMISE MARGINS WHICH MINIMIZE ERROR

3. APPLICATION

RESULTS: TRANSPORTATION MODELING & FORECASTING

1. IDENTIFICATION UNDER LOCAL INCONSISTENCY

PROBLEM: RESOLVING LOCAL INCONSISTENCY AMONG PROJECTIONS

2. CONFORM ALGORITHM

METHOD: OBTAIN COMPROMISE MARGINS WHICH MINIMIZE ERROR

- IRRESOLVABLE LOCAL INCONSISTENCY

- BASIC OPERATIONS OF CONFORM

- SOLUTIONS WITH & WITHOUT CLAMPING

- COMPARISON WITH ARBITRARILY CHOOSING A PROJECTION

- CONFORM EQUATIONS

- CONFORM ALGORITHM

- POST-TEST OF CONFORM RESULTS

3. APPLICATION

RESULTS: TRANSPORTATION MODELING & FORECASTING

1. IDENTIFICATION UNDER LOCAL INCONSISTENCY

PROBLEM: RESOLVING LOCAL INCONSISTENCY AMONG PROJECTIONS

2. CONFORM ALGORITHM

METHOD: OBTAIN COMPROMISE MARGINS WHICH MINIMIZE ERROR

3. APPLICATION

RESULTS: PORTLAND TRANSPORTATION MODELING & FORECASTING

- THREE MAJOR USES OF THIS METHODOLOGY

- TRANSPORTATION MODELING VARIABLES

- SPECIFIC MODELING TASKS

- MODELING EXPERIMENTS

THREE MAJOR USES OF THIS METHODOLOGY

THE GENERAL IDEA

1. **INTEGRATE** DATA FROM DIFFERENT SOURCES,

   EVEN WHEN **CONTRADICTORY**.

   e.g., $AB + BC \rightarrow ABC$ (ABC has structure AB:BC)

   CLAMPING OPTIONAL.

2. **OBTAIN HIGHER ORDER STRUCTURE**

   WITH SUPPLEMENTARY VARIABLES.

   e.g., $ABT + BCT \rightarrow ABCT \nrightarrow ABC$ (full ABC structure)

   CLAMPING OPTIONAL.

3. USE CROSS-TABULATIONS FOR **FORECASTING**.

   E.G., $A\,B + A_f \rightarrow A_f\,B_f \rightarrow B_f.$

   B IS <u>NOT</u> CLAMPED.

TRANSPORTATIONAL MODELING VARIABLES

PORTLAND TRIP GENERATION BASED ON FOLLOWING VARIABLES:

H: HOUSEHOLD SIZE

I: INCOME

A: AGE OF HOUSEHOLDER

Z: GEOGRAPHIC ZONE

SUPPLEMENTARY VARIABLES

T: TENURE (OWN VS. RENT)

      "NUISANCE VARIABLE"

U: # UNITS IN RESIDENCE STRUCTURE

      "DRIVING VARIABLE"

SPECIFIC MODELING TASKS

INVOLVING RESOLUTION OF LOCAL INCONSISTENCIES

1. INPUT: PROJECTIONS of H I A Z   or of H I A Z T

    METHOD: CLAMPING OPTIONAL

    OUTPUT: H I A Z

    USE: # 1 (INTEGRATION) + #2 (HIGHER STRUCTURE)

    PURPOSE: CALIBRATE TRANSPORTATION SIMULATIONS

2. INPUT: H I A Z and A_f, Z_f or A_f, $(ZU)_f$

    METHOD: NO CLAMPING

    OUTPUT $H_f I_f A_f Z_f$, i.e., PREDICT $(H I)_f$

    USE: #3 (FORECASTING)

    PURPOSE: ESTIMATE FUTURE DEMAND

MODELING EXPERIMENTS

DIFFERENT MODELING EXPERIMENTS:

	INPUTS	OUTPUTS	FORECAST VAR.
1	HZ, IZ, AZ	HZ:IZ:AZ	A,Z
2	HZ, IZ, AZ, HA, IA	HZ:IZ:AZ:HA:IA	A,Z
3	HZT, IZT, AZT, HAT, IAT	HIAZ	A,Z
4	HZT, IZTU, AZT, HAT, IAT	HIAZU	A, ZU

ENRICHING STRUCTURE OF HIAZ:

3: HAT:IZT: IAT:HZT: AZT →HIAZT   →HIAZ
4: HAT:IZTU: IAT:HZT: AZT →HIAZTU →HIAZU

THE DIFFERENCE BETWEEN 3 & 4 IS ONLY IN FORECAST

MODEL PERFORMANCE

MODEL	ASSOC. FOR HI, HA, IA	ERROR
1	NO, NO, NO	30
2	NO, YES, YES	17
3, 4	YES, YES, YES	NOT YET AVAILABLE

ERROR (EMPIRICAL) = WORST DIFFERENCE FROM LINE COUNTS