I have given you a copy of a paper ‘Imperfect Price Information and Valuation by Comparable Sales’ by D.G. Wiltshaw in the Journal of Property Research.

The main idea of predicting the value of a house from comparable sales is sound and is done all the time with, of course, some errors. A good property valuer will tell you the confidence they have in their estimates. One would hope with standard properties to have estimates with errors of no more than 10%.

It seems a reasonable first step to identify some key characteristics and then list the sales prices for some comparable listings as a function of the characteristics. A standard mathematical/statistical idea is to attempt to fit a hyperplane through the points and then use this hyperplane to estimate sales prices for the given property to be valued. One could quibble with the assumption of linearity; experience in the real estate field might suggest different functions than hyperplanes but let us focus now on hyperplanes.

1. The main flaw I see in this model is that it assumes that the hyperplane goes through the origin (in the paper the hyperplane is \( \sum_{i=1}^{n} k_i p_i = 0 \)) Thus there is a missing constant. Think of a tangent line (or other line) approximating an arbitrary function locally. Surely you wouldn’t expect the line to go through the origin. If you required the line to go through the origin you’d expect that the line would be a poorer fit.

2. The second main flaw is in asking for equalities. Surely a ‘best’ hyperplane fit is what you are seeking. If nothing else, the sales prices are surely themselves error prone (unless you believe in a perfect market!). But the real reason is that you don’t expect to get a perfect fit to a hyperplane since it is a flawed model anyway.

The equalities are what got the Commerce student in trouble. She used real data and not surprisingly the fit was not perfect.

One can pick holes in all aspects of the model but I think the above two are most important. But here are some others.

1) multiplication. \( k_{ij} p_j \) which is attribute quantity \( \times \) imputed price of attribute. Surely property value is not a linear function of lot size; there is an decreasing value per unit area. This makes no sense in the context of age of a house since old could be charming or age could mean outdated. A quantity of 0 or 1 for a garage is too simplistic; surely a large new double garage is more valuable than a old leaky one car garage.

2) addition. \( \sum_j k_{ij} p_j \). Does it make sense to add these together? I don’t think it make sense as other than a statistical model. One would expect both positive and negative synergy between the characteristics. For example: a small house on a large lot. This is a classic situation in Vancouver where the house gets valued at near zero.

3) equalities. It does not seem reasonable to expect that every comparable sale was made at a perfect price. The state of the buyer and seller (and the market) do affect the price paid. Thus even if you believed the model up to this point, you would probably merely assume that \( \sum_j k_{ij} p_j \) lies in an interval say with 5% of the sales price \( c_i \).

4) maximize. It makes no sense to me to maximize the resulting appraisal. It might make sense to obtain both the maxiumum and minimum and use that as an interval of potential values appraising the property somewhere in the middle of that range.

5) positivity. For me it is not clear from the initial conditions that all imputed prices should be positive. A case in point is the presence of a swimming pool which is usually a negative
(in the Vancouver market). If you view the model as merely an exercise in curve fitting, then you would surely agree that allowing negative imputed prices might improve the accuracy of the fit.

By this point, I have criticized almost everything in the model. I nonetheless believe it possible that a program could be developed which would predict with some accuracy. If nothing else exercises of this sort refine the processes used by appraisers. In my one attempt to get a professional appraisal, the first lines of that appraisal referred to the lack of comparable sales and hence the uncertainty of the appraised value.

Two group of students attempted to use an altered form of this model to predict prices. Despite access to a substantial database of Vancouver sales, they were unable to predict prices with an accuracy of 10%. Thus more work has to be done on this model.