♦ MoHPC ♠   *The Museum of HP Calculators*

HP Forum Archive 18

[Return to Index | Top of Index]

Solving a linear system with physical meaning

*Message #1 Posted by **Rodger Rosenbaum** on 19 Jan 2008, 8:47 p.m.*

On page 18-12 of the HP48G User's Guide, it says:

"Be sure to pay attention to the nature of the linear system you are solving because it will influence how you should interpret the result array. In some cases, you will want to test for ill-conditioning before accepting even an exact solution as a "true" solution."

I will give a system that exemplifies a situation where the exact solution should probably not be considered the "true" solution.

Imagine that you need to measure an electric current, and to increase accuracy you will make the measurement with 4 nominally identical meters and derive some kind of weighted average of the readings of the 4 meters. These meters are just simple analog panel meters.

To get a weighting factor for each meter, you connect the 4 meters in series and pass an accurately known current through all 4 meters at the same time. Pass a current of 1, 2, 3 and 4 amperes through the meters, and note the reading on each meter. Create a matrix of the meter readings and set up a linear sytem like this:

A*W = B

A is the matrix of meter readings and B is the column matrix of the accurate currents that were passed through the meters in series. W will be the weighting factors to be applied to the meter readings in the future when this arrangement is used to measure currents. Each column is the 4 readings from a particular meter for the 4 currents. Each row is the readings from the 4 meters for a particular current.

The matrix A of readings is:

```
[[ 0.95 1.09 1.01 1.09 ]
 [ 1.94 1.95 2.02 1.91 ]
 [ 2.92 3.00 2.93 2.98 ]
 [ 4.01 3.95 3.97 3.94 ]]
```

The column matrix of currents is:

```
[[ 1 ]
 [ 2 ]
 [ 3 ]
 [ 4 ]]
```

What we want to do is get a vector W which will contain 4 numbers, weighting factors to be applied to the meter readings. In other words, if the 4 meter readings are m1, m2, m3 and m4, and the weighting factors (elements of W) are w1, w2, w3 and w4, then the derived current value will be w1*m1+w2*m2+w3*m3+w4*m4. Obviously, given that the 4 meters are nominally identical, weighting factors of w1=.25, w2=.25, w3=.25 and w4=.25 are close to what we want. Since the meters will have some unavoidable manufacturing tolerances, the weighting factors will be slightly perturbed from the ideal values I just gave.

Other solutions to the problem would be W vectors of (the T means transpose):

[1 0 0 0]T, just use the readings from the 1st meter.

[0 1 0 0]T, just use the readings from the 2nd meter, etc.

[.5 .5 0 0]T, just use the readings of the 1st and 2nd meters.

[.5 0 .5 0]T, just use the readings of the 1st and 3rd meters, etc.

If we solve the A*W = B linear system on the HP48G (or its descendants), we get:

```
[[  .8822598871   ]
 [  3.1448587566  ]
 [  -.4836158193  ]
 [ -2.5482485873 ]]
```

While this is an exact mathematical solution, it doesn't make sense given the physical situation. Two of the weights are negative! And the system isn't particularly ill-conditioned; the condition number of the A matrix is about 1000, so we lose about 3 digits of accuracy in the solution. What we would like is 4 weights of about .25 each.

Here is a case where the exact mathematical solution doesn't make sense. It reminds me of those freshman physics problems where you get two solutions to a problem, one of which doesn't fit the physical situation.

What shall we do?

Re: Solving a linear system with physical meaning
Message #2 Posted by John Keith on 20 Jan 2008, 3:07 p.m.,
in response to message #1 by Rodger Rosenbaum

Hi Roger,

I'm no expert at linear algebra, but I can see a couple of possible ideas. Basically, I'm not sure that a system of linear equations is the right model for this problem, which may be why the answers don't make sense.

The inaccuracies in electric meters (and most other things in the real world) are non-linear, and the values in your matrix A bear this out. If you look at the lowest and highest reading for each row or column, they're all over the place. Which means that if they represent data for real meters, those meters have large non-linearities which would need 3rd order (at least) polynomials to model them.

Furthermore, since neither meter reads consistently low nor high, there is probably no single "correction factor" that could be applied to a given meter that would be accurate for all readings. The best you could do is a compromise, which is indeed why one makes such measurements at multiple current levels.

Mathematically speaking, this sounds to me like a minimization problem- a search for a best overall solution, rather than an exact solution. Perhaps linear programming or some non-linear optimization algorithm would be best, and hopefully others with more knowledge can be of help here.

John

Re: Solving a linear system with physical meaning
Message #3 Posted by Rodger Rosenbaum on 20 Jan 2008, 4:55 p.m.,
in response to message #2 by John Keith

Just to clarify, the data aren't real measurements. I started out with an ideal matrix and added random perturbations. Any trend is accidental, an artifact. So any method of solution intended to be general shouldn't make special provisions for apparent trends in these particular data.

This whole thing is intended to be illustrative of the HP48G user manual advice that sometimes the exact solution might not be what you want.

The question is, how SHOULD we handle the problem?

Re: Solving a linear system with physical meaning
Message #4 Posted by Jonathan Eisch on 20 Jan 2008, 5:13 p.m.,
in response to message #3 by Rodger Rosenbaum

It seems to me that if you have four meters, with random errors (uncorrelated with any known quantity), any combination of these meters (with weights equaling one, of course) are equally valid. Garbage in, garbage out.

In other words: Why would you expect [.25 .25 .25 .25] when [1000 -999 0 0] is just as good?

It seems to me that you would have to minimize something to deal with your random errors.

If your errors were linear with respect to current, then I would guess that your system makes some sense. Why not try that?

-Jonathan

Re: Solving a linear system with physical meaning

Message #5 Posted by Rodger Rosenbaum on 20 Jan 2008, 9:14 p.m.,
in response to message #4 by Jonathan Eisch

Quote:

It seems to me that if you have four meters, with random errors (uncorrelated with any known quantity), any combination of these meters (with weights equaling one, of course) are equally valid. Garbage in, garbage out.

In other words: Why would you expect [.25 .25 .25 .25] when [1000 -999 0 0] is just as good?

It seems to me that you would have to minimize something to deal with your random errors.

If your errors were linear with respect to current, then I would guess that your system makes some sense. Why not try that?

-Jonathan

The idea is that if the errors are zero mean, uncorrelated with each other, then if you include more meter readings, the desired signal (actual value of the current) is accumulated while the errors cancel each other, in a manner of speaking. That is why [.25 .25 .25 .25]T is better. "Minimizing something" is as easy as just adding up weighted readings from a bunch of meters, the more meters the better.

My point here is to show that the standard way of solving a linear system doesn't give a good result in a case like this.

Imagine if the A matrix were:

```
[[ 1 1 1 1 ]
 [ 2 2 2 2 ]
 [ 3 3 3 3 ]
 [ 4 4 4 4 ]]
```

If you try to solve the system in the usual way, you get an error. But, plainly, the solution [.25 .25 .25 .25]T is the desired one. (There an infinite number of "exact" solutions in this case.)

And, in fact, solving this system with the LSQ function on the HP50 gives just that result. But, it won't work with the A matrix I gave in my first post in this thread.

Re: Solving a linear system with physical meaning

Message #6 Posted by Dave Shaffer (Arizona) on 20 Jan 2008, 9:58 p.m.,
in response to message #5 by Rodger Rosenbaum

This is NOT a linear system - as John pointed out, your meter readings are non linear (and, I suspect, if you made a plot also have non-zero intercepts, which I think you have not properly accounted for either).

So, certainly the manual's warning is more than a propos about blinddly accepting a "solution"!

If you found least squares fits for the behavior of each meter, along with some error estimates, you could probably formulate a much better and more legitimate averaging routine.

Re: Solving a linear system with physical meaning

Message #7 Posted by Rodger Rosenbaum on 21 Jan 2008, 4:40 a.m.,
in response to message #6 by Dave Shaffer (Arizona)

> Quote:
>
> This is NOT a linear system - as John pointed out, your meter readings are non linear (and, I suspect, if you made a plot also have non-zero intercepts, which I think you have not properly accounted for either).
>
> So, certainly the manual's warning is more than a propos about blinddly accepting a "solution"!
>
> If you found least squares fits for the behavior of each meter, along with some error estimates, you could probably formulate a much better and more legitimate averaging routine.

As I mentioned in the first post, we're trying to solve w1*m1+w2*m2+w3*m3+w4*m4 = b1, etc. Like many a least squares problem, the problem is linear in the weights, although this is not a least squares problem since we have 4 equations and 4 unknowns.

The meter readings are not non-linear; they are linear with the addition of noise. That is the sort of situation where averaging of some sort can reduce the effect of the noise.

Re: Solving a linear system with physical meaning

Message #8 Posted by Jonathan Eisch on 20 Jan 2008, 11:09 p.m.,
in response to message #5 by Rodger Rosenbaum

> Quote:
>
> My point here is to show that the standard way of solving a linear system doesn't give a good result in a case like this.

No, you got exactly what you asked for. (A great result!) You wanted some set of weights to make the readings you provided equal the values you provided. If you do the matrix multiplication with the weights that you got, you'll see that you get 1,2,3,4 to within the precision you provided. That looks like success to me.

If you do the matrix multiplication with the .25,.25,.25,.25 matrix you want, you get something like 1.035, 1.955, 2.9575, 3.9675 with is clearly further away, so clearly not a mathematically better set of weights.

> Quote:
>
> Imagine if the A matrix were:
>
> ```
> [[ 1 1 1 1 ]
>  [ 2 2 2 2 ]
>  [ 3 3 3 3 ]
>  [ 4 4 4 4 ]]
> ```
>
> If you try to solve the system in the usual way, you get an error. But, plainly, the solution [.25 .25 .25 .25]T is the desired one. (There an infinite number of "exact" solutions in this case.)

Mathematically you'd be minimizing a flat parameter space, nothing matters, so no solution would be mathematically preferable over any other. In other words, numbers don't /desire/ anything. ;-)

But you didn't use this perfect A matrix, you used a slightly different one, so now your flat space has tiny bumps, and your minimization rolled into one of them. Try giving it a (linear) slope and see what you get.

-Jonathan

Re: Solving a linear system with physical meaning

Message #9 Posted by Rodger Rosenbaum on 21 Jan 2008, 5:55 a.m.,
in response to message #8 by Jonathan Eisch

> Quote:
>
> No, you got exactly what you asked for. (A great result!) You wanted some set of weights to make the readings you provided equal the values you provided. If you do the matrix multiplication with the weights that you got, you'll see that you get 1,2,3,4 to within the precision you provided. That looks like success to me.

I failed to be sufficiently explicit in my first post in describing what this is all about. The idea here is to find some weights with the "calibration" measurments I described in the first. Then subsequent measurements would be used on another day to measure some currents, and the derived weights would be applied to those subsequent measurements. The current(s) to be measured would must likely be different than the initial "calibration" currents, and would not get the nearly exact (to about 10 digits) results you get if you apply the weights:

```
[[  .8822598871   ]
 [ 3.1448587566   ]
 [ -.4836158193   ]
 [ -2.5482485873 ]]
```

To be more concrete, let's suppose I make some more measurements to see how well the weights we determined in the first set of measurements work with some new current measurements. The ideal currents applied this time will be:

```
[[ 1.5 ]
 [ 2.5 ]
 [ 3.5 ]
 [ 4.5 ]]
```

and the measured currents might be (perturbed from the exact values with random noise):

```
[[1.44 1.42 1.41 1.54 ]
 [ 2.52 2.41 2.42 2.47 ]
 [ 3.58 3.43 3.43 3.55 ]
 [ 4.59 4.41 4.58 4.59 ]]
```

The weights calculated with the standard linear system solver are:

```
[[ -5.8890631686 ]
 [ 10.0895955135 ]
 [ -.02138884926 ]
 [ -2.803127525 ]]
```

If we apply these weights to the second set of measurements just above, we get about 10 digit agreement with the applied currents. But, if we apply the weights derived in the first post, namely:

```
[[ .8822598871   ]
 [ 3.1448587566  ]
 [ -.4836158193  ]
 [ -2.5482485873 ]]
```

we get (rounded to 6 digits):

```
[[ 1.12995 ]
 [ 2.33788 ]
 [ 3.24027 ]
 [ 4.00698 ]]
```

If we just use the intuitively reasonable weights [.25 .25 .25 .25]T, we get:

```
[[ 1.4525 ]
 [ 2.455  ]
 [ 3.4975 ]
 [ 4.5425 ]]
```

considerably better. And perhaps we can do even better.

The idea is to find some weights such that if applied to current readings contaminated by noise (errors), the result won't be perturbed much away from the true value. Using the weights derived from an exact linear system solution is not a good thing; these weights will magnify the noise by an amount about equal to the condition number of the A matrix used to derive the weights.

The reason it isn't good is because this is a system where the columns of the A matrix are almost carrying the same information (almost linearly dependent). This means that the condition number of the matrix is likely to be large. An exact solution of the system works well for the A matrix used in the solution, but another matrix with different perturbations (additive noise) may not give a reasonable solution. One way to deal with this (sometimes recommended for ill-conditioned systems) is to reduce the order of the system. Just delete the columns that are linearly dependent on the others. In this case, just select *one* of the columns. But, even though the columns are nearly linearly dependent, there may be

some information to be gained if all of them are included. The problem is how to do this without getting the error magnification that using standard linear system solution methods gives.

The physical nature of this particular problem shows that weights of [.25 .25 .25 .25]T are probably close to the ones we want. --

I said:

> Quote:
> _____
>
> Imagine if the A matrix were:
>
> ```
> [[ 1 1 1 1 ]
>  [ 2 2 2 2 ]
>  [ 3 3 3 3 ]
>  [ 4 4 4 4 ]]
> ```
>
> If you try to solve the system in the usual way, you get an error. But, plainly, the solution [.25 .25 .25 .25]T is the desired one. (There an infinite number of "exact" solutions in this case.)
>
> _____

Then you said:

> Quote:
> _____
>
> Mathematically you'd be minimizing a flat parameter space, nothing matters, so no solution would be mathematically preferable over any other. In other words, numbers don't /desire/ anything. ;-)
>
> _____

Look again at the subject line. It's not true that "nothing matters" given the physical nature of the setup. The reason for preferring the [.25 .25 .25 .25]T weights is not mathematical, or because the numbers "desire" something. When I said "desired" solution, I was referring to *my* desire, not the desire of the numbers. The reason is the physical situation. A person could just select one of the meter readings and just ignore the rest, but then you would be throwing away possible information carried by the other readings. If you use the [.25 .25 .25 .25]T weights, you are averaging the readings and presumably reducing the effect of the noise. Because I chose to use identical meters in this hypothetical situation, it's easy to see approximately what the weights should be. Similar situations may arise where it isn't so obvious. What mathematical procedure can we follow to get a good set of weights, less affected by the ill-conditioning of the A matrix?

Re: Solving a linear system with physical meaning

*Message #10 Posted by **Jonathan Eisch** on 21 Jan 2008, 11:36 a.m.,*
in response to message #9 by Rodger Rosenbaum

> Quote:
>
> What mathematical procedure can we follow to get a good set of weights, less affected by the ill-conditioning of the A matrix?

Averaging?

-Jonathan

Re: Solving a linear system with physical meaning

*Message #11 Posted by **Nick** on 21 Jan 2008, 7:59 p.m.,*
in response to message #9 by Rodger Rosenbaum

Rodger, this doesn't have to do with mathematically exact solutions that aren't possible physically, like for example partcles with imaginary mass fulfilling special relativity and travelling at speeds greater than c. The result that you got is:

1) Absolutely valid for itself, since there is absolutely no reason for weights to be always positive. Even absolute temperature is "formally" accepted to be "negative" (i.e. under 0K) when used only as a "weight" for entropy differences.

2) Even if we want all weights to be positive, your underlying mathematical model doesn't include this as an additional constraint, and hierin lies the problem, as Jonathan also described very nicely with "bumps". You know that all weights have to be the same, but your model doesn't and so it goes down the bumps. You must explicitely formulate the constraint that w1=w2=w3=w4=1/n where n is the number of your current meters, or anything else in a repeated measurement. You do statistics with a constraint here, but you don't formulate the constraint explicitely - or using other words you want a flat world but your mathematical model has bumps and no breaks to prevent the "moving things" (weights) to fall into the bumps as they search for their "best" values. (The word "best" means arithetically best here.

So it is not a subject of mathematical solutions being physically impossible - it is a matter of inadequate mathematical description of the system iff we demand all weights to be positive, and even more so iff we demand them to be all positive and all equal to each other.

I try to present an idea here using as few words as possible and not having any possibility (or knowledge ;-)) for using textbook notation for formulae.

Let's reduce our sight to one of the current meterings, say A1 (1 Ampere). In the ideal case of ideal meters all meterings would be A1=A=1, the "real" current. So the first row of A would be [1 1 1 1] in its ideal version, and the corresponding ideal element of B would be also 1. Then, for the first row of the matrix equation we have:

$SUM(A1i*Wi, i=1..n)=B1$

in the most general case, and:

$SUM(A1*Wi, i=1..n)=B1$ in the special "ideal" case. Here we are in the flat world that Jonathan mentioned. Infinite many solutions as you also mentioned and even if you know about some of them, that "feels" better, we didn't told our system about our feelings yet, since the constraint is missing! Iff for all Wi it is valid: $Wi=W=1/n$, then the equation for the first row goes to:

$SUM(A1*W, i=1..n)=B1 => SUM(A1, i=1..n)*W=B1 => n*A1*W=B1$

which for $n=4$, $A1=1$, $B1=1$ has one solution $W=1/4$. Notice that we get that solution, only *after* introduction of the contstraint!

Let's move on to the non-ideal case, where A11, A12, A13, A14 are different. We can formulate each of them as the "real" A1 plus some Delta that can be positive or negative. Denote that with DA11, DA21, and so on. Then for the first row again:

$SUM(A1i*Wi, i=1..n)=B1$

Which means that:

$SUM((A1+DA1i)*Wi, i=1..n) = B1 =>$

$SUM(A1*Wi, i=1..n)+SUM(DA1i*Wi, i=1..n)=B1$

This world has bumps and the Wi's will fall in there if we not explicitely specify again for all Wi: $Wi=W=1/n$. After this constraint we have:

$SUM(A1*W, i=1..n)+SUM(DA1i*W, i=1..n)=B1 =>$

$SUM(A1, i=1..n)*W+SUM(DA1i, i=1..n)*W=B1 =>$

$n*A1*W+SUM(DA1i, i=1..n)*W=B1 =>$

$(n*A1+SUM(DA1i, i=1..n))*W=B1$

which again will have a solution of somewhere around $W=1/n$ since the sum of all Deltas, $SUM(DA1i, i=1..n)$, is about zero. Again we get this result *after* introduction of the constraint!

Putting all As and Ws and Bs into a matrix equation is much the same, but still the constraint *has* to be specified explicitely, in order to have an adequate mathematical description of what we want. Actually the matrix equation goes over to vector equations then, where the constant W can be calculated to 1/n (or 1/4 in your specific case).

Nonetheless, as already said, there is also absolutely nothing wrong with negative weights, should the non-restricted model be considered good enough.

Thanks a lot for this very interesting problem - a very good demonstration of how some "intuitive feeling" can slip into mathematics and cause surprises if the formalism is not describing exactly what it was intended to desccribe. As Russel once said: The greatest challenge to any thinker is stating the problem in a way that will allow a solution. ;-)

But then again, all we intuitively considered true proved to be false! (From Russel too?) ;-)

Cheers and excuse my lust for maths late in the night!

Nick

Re: Solving a linear system with physical meaning

Message #12 Posted by Rodger Rosenbaum on 22 Jan 2008, 5:11 a.m.,
in response to message #11 by Nick

> Quote:
>
> ---
>
> Rodger, this doesn't have to do with mathematically exact solutions that aren't possible physically, like for example partcles with imaginary mass fulfilling special relativity and travelling at speeds greater than c.
>
> ---

It may not have to do with "particles with imaginary mass", but it most certainly has to do with mathematically exact solutions being unsatisfactory for reasons having to do with the physical aspects of the problem. That is what I intended when I started this thread, and I hinted at that in the subject line.

Let me give a concrete example that I also hinted at in the first post. Consider this freshman physics problem.

From a point on a platform, 10 feet above the ground, launch a steel ball upward with an initial velocity of 60 feet per second (perhaps at a slight angle so it doesn't fall back on your head!). How many seconds does it take until the ball reaches the ground? Taking the acceleration of gravity as 32.17 f/s, we set up the following equation:

$16.085*t^2 - 60*t - 10 = 0$

Solving, we get two solutions, t=3.89 and t=-.1598

Do we really believe that the ball will hit the ground -.1598 seconds after we launch it, and again in 3.89 seconds?

The problem I have given in this thread is similar in that there are numerous solutions that seem reasonable along with the one mathematically exact solution. We must look to the physical nature of the problem to help us select the appropriate solution.

Let me digress.

In the November 1970 issue of The American Mathematical Monthly, there is an article titled "Pitfalls in Computation, or Why a Math Book isn't Enough", by Professor G. E. Forsythe. He gives an example:

Here is a small linear system, A*x = B:

```
[[ .780 .563 ] * [[ x1 ]  = [[ .217 ]
 [ .913 .659 ]]   [ x2 ]]    [ .254 ]]
```

Two solutions to the system are proposed, [.999 -1.001]T and [.341 -.087]T.

Which one is better? The usual check is to substitute them both into the original problem. Substituting [.999 -1.001]T gives a residual (A*x-B) of [-.001343 -.001572]T. Substituting [.341 -.087]T gives a residual of [-.000001 0.0]T. It seems clear that the second proposed solution is better than the first, since it makes the *residual* far smaller.

However, in fact the true solution is [1 -1]T as the reader can easily verify. Hence, the first proposed solution is far closer to the true solution than the second.

A persistent person may ask again: which solution is *really* better? Clearly, the answer must depend on one's criterion of goodness: a small residual, closeness to the (mathematically) true solution, or something else. Surely one will want different criteria for different problems.

The problem I gave has physical aspects that lead to certain choices for a solution. First of all, the meters are nearly identical. That leads to the notion that they should be weighted nearly equally. Also, I assumed that the noise (errors) is zero mean, uncorrelated between meters. This is a reasonable assumption, absent any evidence to the contrary. In a real situation, there might be reason to assome otherwise. If so, then it will have to be dealt with some other way, but for the purposes of this thread, that's the assumption.

Because the noise tends to be reduced by averaging processes, whereas the desired quantity (value of the current) isn't, we don't want to do what Jonathan Eisch suggested:

Quote:

In other words: Why would you expect [.25 .25 .25 .25] when [1000 -999 0 0] is just as good?

The reason the weights [1000 -999 0 0] are not just as good is because those weights will greatly magnify the errors with respect to the desired current reading. Plus, of course, the readings from two of the meters are thrown away, losing any helpful information they may carry. The weights are mathematically equivalent if the current readings are identical for each meter, but that's not the case here.

And besides the purely physical consideration that we would like to include the readings from all 4 meters (otherwise why bother having 4 meters?), there is a mathematical reason:

In this problem, the sum of the weights in any of the solutions we've seen so far is nearly 1. It is well known that, given 4 numbers whose sum is 1, the root-mean-square value of the 4 is a minimum when they are all 4 equal. Since the additive noise in the readings is uncorrelated, a weighted sum of the 4 noise signals will be minimized when the weights are nearly equal.

So, let's assume that we have determined that the weights should be *nearly* [.25 .25 .25 .25]T, which are far from the weights the exact linear system solution gives.

It would appear that our criterion should be to judge the *goodness* of our solution by looking at the norm of the residuals. Calculate this on the HP50 by using the ABS function applied to (A*w-B).

Let's see if we can find the minimum residual norm solution where the weights are constrained to be equal. First so we will have a basis for comparison, let's compute the residual norm with weights of [.25 .25 .25 .25]T (using the A matrix I gave in the first post of this thread). I get .07818.

To calculate the residual norm where the weights are constrained to be equal, multply A by [1 1 1 1]T and get [4.14 7.82 11.83 15.87]T. Put this on level 1 of the stack and put [1 2 3 4]T on level 2 of the stack; press LSQ and get [[.252607010711]]. Using [.252607 .252607 .252607 .252607 }T as a solution vector, compute the residual norm. I get .0540137, a substantial improvement over .07818.

It is possible to do even better if the individual weights are not constrained to be equal.

So far, nobody has given much in the way of numerical solutions. There have been discussions of various procedures, but no examples to show if they give a "good" result, with "good" meaning the sort of thing I've discussed.

Nick said:

Quote:

Nonetheless, as already said, there is also absolutely nothing wrong with negative weights, should the non-restricted model be considered good enough.

There are several things wrong with negative weights, and I hope I've explained why.

Now, please use your calculators to find some weights with even lower residual norm, but close to [.25 .25 .25 .25]T, and tell us how you did it.

But, be aware that there is a trade off. If you do a search for a global minimum in the residual norm, without constraints the search takes you to the exact linear system solution, with a residual norm of zero. It would seem that if you allow the individual weights to vary, you should be able to get a smaller residual norm, with the norm getting smaller as you approach the exact linear system solution.

What I'm looking for from all you who are interested is a procedure to find a set of weights with better residual norm than the one where the weights are constrained to be equal. And, try to do it with linear algebra techniques.

Re: Solving a linear system with physical meaning

Message #13 Posted by Stefan K. on 23 Jan 2008, 2:42 a.m.,
in response to message #3 by Rodger Rosenbaum

Quote:

The question is, how SHOULD we handle the problem?

Just to add my 2 cents (and summarizing what others in this thread alrady have pointed out):

You try to make **non-exact** measurements (due to non-linearity of meter, reading and random errors...) **exact** by "calibrating" a set of 4 meters at 4 points (4 different currents) - that's impossible.
Of course you can do it the way you did it, but it's neither stable nor does it make sense (negative coefficients).

The way I would do it is linear regression: you try to find a linear function (y=k*x+d) with unknown coefficients k and d, and fit them by minimizing the error (sum(current-y)^2). Actually you end up with a set of linear equation agains (for k and d), which you can solve by inverting a matrix (but since it's 2x2, you can almost do it in your head).
For setting up the matrix and disturbance vector, have a look at Linear regression at wikipedia .

Stefan

Re: Solving a linear system with physical meaning

*Message #14 Posted by **Rodger Rosenbaum** on 23 Jan 2008, 3:49 a.m.,*
in response to message #13 by Stefan K.

Stefan, you're putting words in my mouth. Look at the beginning of my first post. I said:

"Imagine that you need to measure an electric current, and to **increase** accuracy you will make the measurement with 4 nominally identical meters and derive some kind of weighted average of the readings of the 4 meters."

I did **not** say, and I'm quoting you:

> Quote:
>
> ...make **non-exact** measurements (due to non-linearity of meter, reading and random errors...) **exact** by "calibrating" a set of 4 meters at 4 points (4 different currents)...

I never claimed to make non-exact measurements become exact.

You said:

> Quote:
>
> Of course you can do it the way you did it, but it's neither stable nor does it make sense (negative coefficients).

Exactly what are you referring to when you say "...the way you did it...". Do you mean the weights derived from a standard linear system solution? I only showed those weights as a bad set of weights. I specifically did not advocate their use. If you think you're rebutting my advocacy of those weights, your rebuttal is misplaced, because I don't advocate them.

And again, from my very first post:

"While this is an exact mathematical solution, it doesn't make sense given the physical situation. Two of the weights are negative!"

Perhaps you should re-read my first post.

> Quote:

The way I would do it is linear regression: you try to find a linear function (y=k*x+d) with unknown coefficients k and d, and fit them by minimizing the error (sum(current-y)^2). Actually you end up with a set of linear equation agains (for k and d), which you can solve by inverting a matrix (but since it's 2x2, you can almost do it in your head).

Why don't you do what Valentin did and what I did: work out an example and show us some numbers. Show that your technique would have a smaller residual norm than Valentin's weights, for example.

Re: Solving a linear system with physical meaning
Message #15 Posted by Stefan K. on 23 Jan 2008, 5:10 p.m.,
in response to message #14 by Rodger Rosenbaum

Hi,

sorry that I didn't understand and appreciate the full depth of the problem initially. I did some thinking (it's a real pity I don't get to do this for work anymore), and I came up with a solution.

First to clarify some points:

As the other Stefan already pointed out, only the same number of readings and meters makes your solution look well defined. Choose 3 readings and you end up with an underdetermined set of equations, choose 5 or more and you get an overdetermined one.

If I inspect your set of equations, A*W=B, the way I interpret it is:

B are the real (ie. error-free) values, I wish I had an instrument in my lab that would give me these, then I could throw away the other four ;-) . A are the readings of the erroneous instruments, and you try to find weighting factors W, that give you the **real** values by a clever combination of the erroneous ones. That's what I meant that you calculate the exact values from non-exact ones. If your instruments only have systematic errors, this would be possible, if you add random errors, this would only work by chance. And it's only possible because you have as many readings as meters. Furthermore, you only know that you get the real values only for the points you already have got - there is no big information gain, since you can't be sure that you get complete nonsense if you interpolate.

But since we already agreed that this is no good solution, here is how I would do it:

First define what B really is: it's some control parameter, it could be what you think the real current is, or just your voltage source value divided by 1 Ohm... Then you need a model, for simplicity let's assume that your current reading is linearly dependent on the control parameter,
A=k*B+d
(with four readings, you could assume a polynomial up to third order). The still unknown parameters k and d will differ for every instrument. In our case k should be close to 1, and d close to 0. In order to determine them, we use linear regression:
\chi^2=\sum (A_i-k*B_i-d)^2
is the error between measurement and linear model (index i is for # of measurement). The minimum of the error is found by solving the equation

```
[[B_i^2  B_i]      [[k]      [[A_i*B_i]
                *          =
 [B_i    N  ]]      [d]]       [A_i]]
```

(you have to do the sums over the parameters with the index, N=4 is number of measurements).
Once you have the coefficients for the polynomial, you can also calculate the error \chi^2 for every instrument, and use this as measure for your weighting factors. I suggest to use
1/(1+\chi^2)
(I added 1 to the denominator in case the error becomes accidently 0, it could be also some other number), of course you have to normalize them so that they add up to 1. But since \chi^2 is always positive, you won't get unreasonable weighting factors. And this also gives you a sound way to get interpolated values.

I try to do a sample calculation tomorrow (after some sleep and when I have access to matlab again).
Stefan

Re: Solving a linear system with physical meaning
Message #16 Posted by Karl Schneider on 22 Jan 2008, 3:21 a.m.,
in response to message #1 by Rodger Rosenbaum

Hi, Rodger --

Interesting problem -- one which I'd never considered.

The determinant of "A" (the matrix of random-error-perturbed measurements) is only 0.0011, making it near-singular. However, a matrix of four *identical* measurements from each of the four meters <u>would</u> be singular. Clearly, the linear solution of this exactly-determined system is the wrong approach for specifying the weights.

This linear-algebra solution, of course, is the one that enables each of the four data points to be matched <u>exactly</u>, but as you showed in a subsequent example, it produced poor results for a different set of quasi-random data.

Since the measurement error is random, calibration of the meters should not be expected to provide any benefit, and it's likely impossible to mathematically derive an "optimal" solution for all expected measurements, based upon only the four provided. Each of the meters has comparable accuracy, so setting each weight to 0.25 is the reasonable thing to do.

As a mathematical exercise, you might try generating additional "measurements" perturbed by an equal degree of random noise for each meter, and see how well the HP-48G's least-squares solution for over-determined systems works. I'd bet that as more measurements are represented in the system, all the weights approach 0.25.

Regards,

--KS

(Edited to sharpen one statement.)

Edited: 22 Jan 2008, 11:45 p.m. after one or more responses were posted

Re: Solving a linear system with physical meaning
Message #17 Posted by Walter B on 22 Jan 2008, 6:07 a.m.,
in response to message #16 by Karl Schneider

Hi,Rodger,

let me second Karl and add my own experimental 20 Milli-Euros : A single measurement means *nothing* unless you *know* the precision of your measuring instrument(s).

In your starting case, you did assume a random measuring error for each instrument, which is perfectly correct. What can be calibrated (physically or mathematically as you want), however, is a systematic (= non-random) deviation between your instruments. This deviation you will only find by repeated measurements, since by doing this you can reduce the influence of the random factors. So, approaching the limit of an infinite number of measurements, your random errors will cancel and only the systematic deviations will survive. While you certainly don't want to measure infinite times, you may get reasonable estimates of the systematic deviations already after 5 measurements per instrument. After all, the basic law of statistics rules: With no information in you'll get no information out ;)

Best regards, Walter

Re: Solving a linear system with physical meaning

Message #18 Posted by Rodger Rosenbaum on 22 Jan 2008, 7:38 a.m.,
in response to message #17 by Walter B

Your comments are certainly something that a person would want to take into account, but they go beyond my point. I was just trying to point out that the exact mathematical solution to the simple situation I posed is not the "best" solution, and is less good the more the meters are identical.

Re: Solving a linear system with physical meaning

Message #19 Posted by Walter B on 22 Jan 2008, 12:40 p.m.,
in response to message #18 by Rodger Rosenbaum

Please consult the recent messages of Valentin and Stefan. Their posts will help you to find the point.

Re: Solving a linear system with physical meaning

Message #20 Posted by Valentin Albillo on 22 Jan 2008, 2:06 p.m.,
in response to message #19 by Walter B

Hi, Walter B.:

Walter B posted:

"Please consult the recent messages of Valentin and Stefan."

Not so *recent*. My message was posted at 6:13 am, well before Rodger replied to you (7:38 am) and Karl Schneider (7:30 am), so he could've read and replied to my message as well as to yours, but he didn't because he hates me and has my messages filtered out :-(

A puny joke, of course ! :-) Thanks for your interest and

Best regards from V.

Re: Solving a linear system with physical meaning

Message #21 Posted by Stefan Vorkoetter on 22 Jan 2008, 4:08 p.m.,
in response to message #18 by Rodger Rosenbaum

Quote:

I was just trying to point out that the exact mathematical solution to the simple situation I posed is not the "best" solution, and is less good the more the meters are identical.

The point is that you've computed the exact mathematical solution to the wrong problem, as Valentin points out below.

Stefan

Re: Solving a linear system with physical meaning

*Message #22 Posted by **Rodger Rosenbaum** on 22 Jan 2008, 7:30 a.m.,*
in response to message #16 by Karl Schneider

Hi, Karl. I was wondering if I would hear from you on this one! I'm also waiting to hear from Palmer.

> Quote:
>
> ---
>
> As a mathematical exercise, you might try generating additional "measurements" perturbed by an equal degree of random noise for each meter, and see how well the HP-48G's least-squares solution for over-determined systems works. I'd bet that as more measurements are represented in the system, all the weights approach 0.25.
>
> ---

Sure enough. You can see that same thing happening if you just perturb the 4x4 A matrix enough times. I got my example matrix by starting with an ideal matrix, adding perturbations of 1%, and computing the condition number. I stopped when I got one with a medium high condition number. Most of the time the condition number's not that high, and the weights computed with a standard linear system solution are in the neighborhood of .25. I wanted a medium high condition number to show what would happen if the meters were nearly identical.

Re: Solving a linear system with physical meaning

*Message #23 Posted by **Valentin Albillo** on 22 Jan 2008, 6:13 a.m.,*
in response to message #1 by Rodger Rosenbaum

Hi, Rodger:

First of all, ***Happy New Year 2008 to everyone***. And now ...

Roger posted:

> *"What shall we do ?"*

Include in your model <u>all known facts</u> , that's what you should do.

There's nothing in your mathematical model that accounts for the empirical fact that you know for sure the meters to be *"nominally identical"* , so given your readings corrupted with random errors, *any* set of weights is mathematically and logically as acceptable as any other.

If you want your mathematical model to agree with your intuition you need to include in it all the known facts, namely the known fact that the meters are *nominally identical*. This means including 6 additional equations, as there are 6 pairings of each meter with one of the others, and for each pairing you expect nominal identity,

This will result in an *overdetermined* linear system of 10 equations in 4 unknowns, which can be readily solved by least squares to determine the 4 weights.

To that effect, I've just concocted the following 3-liner for the HP-71B:

```
10 DESTROY ALL @ OPTION BASE 1 @ INPUT "#E,U=";M,N
20 DIM A(M,N),C(M),B(N,N),D(N),X(N) @ MAT INPUT A,C
30 MAT B=TRN(A)*A @ MAT D=TRN(A)*C @ MAT X=SYS(B,D) @ MAT DISP X;
```

which asks for the number of equations (E) and the number of unknowns (U), then lets you input the known data and proceeds to print out the computed values for the unknowns using least squares methodology.

If applied to your initial, non-overdetermined system, it promptly gives:

```
 >RUN


  #E,U=4,4


  A(1,1)?    .95, 1.09, 1.01, 1.09
  A(2,1)?   1.94, 1.95, 2.02, 1.91
  A(3,1)?   2.92, 3,    2.93, 2.98
  A(4,1)?   4.01, 3.95, 3.97, 3.94


C(1)? 1,2,3,4


    .882259888123
   3.14485876301
   -.483615821069
  -2.5482485928
```

which are the exact weights you found so unexpectedly (give or take a few ulps). Now, if you run it again with the 10-equation system which includes your knowledge of every pair of meters being nominally identical, you'll get

```
>RUN


#E,U=10,4


A(1,1)?     .95, 1.09, 1.01, 1.09
A(2,1)?   1.94, 1.95, 2.02, 1.91
A(3,1)?   2.92, 3,    2.93, 2.98
A(4,1)?   4.01, 3.95, 3.97, 3.94
A(5,1)?   1,  0,  0, -1          1st meter nominally identical to 4th
A(6,1)?   1,  0, -1,  0          1st meter nominally identical to 3rd
A(7,1)?   1, -1,  0,  0          ...
A(8,1)?   0,  1,  0, -1          ...
A(9,1)?   0,  1, -1,  0          ...
A(10,1)?  0,  0,  1, -1          3rd meter nominally identical to 4th


C(1)? 1,2,3,4,0,0,0,0,0,0


.25328021694
.252112563361
.253205061818
.251830755168
```

which are the kind of weights you were expecting to begin with.

The moral:

Thou shalt always include in your mathematical models all the relevant knowledge, lest you'll be surprised when thy results don't agree with thy expectations.

Best regards from V.

Re: Solving a linear system with physical meaning

Message #24 Posted by Antoine M. Couëtte on 22 Jan 2008, 8:11 a.m.,
in response to message #23 by Valentin Albillo

Wonderful, Valentin !!! You're back !! :-) + -:))

HAPPY NEW YEAR TO YOU TOO

Antoine

Antoine M. Couëtte

Edited: 22 Jan 2008, 8:12 a.m.

Why, thank you very much, Antoine ! :-) [NT]
Message #25 Posted by Valentin Albillo on 22 Jan 2008, 8:23 a.m.,
in response to message #24 by Antoine M. Couëtte

Best regards from V.

Re: Why, thank you very much, Antoine ! :-) [NT]
Message #26 Posted by Antoine M. Couëtte on 22 Jan 2008, 9:15 a.m.,
in response to message #25 by Valentin Albillo

... simply because I had not read from you on this forum for quite some time (a few moons already) ...

Antoine

Re: Solving a linear system with physical meaning
Message #27 Posted by Antonio Maschio (Italy) on 22 Jan 2008, 10:19 a.m.,
in response to message #23 by Valentin Albillo

Well, if you don't mind, could you explain to me what's the purpose of the SYS() instruction?

Taken from your program:

Quote:

```
10 DESTROY ALL @ OPTION BASE 1 @ INPUT "#E,U=";M,N
20 DIM A(M,N),C(M),B(N,N),D(N),X(N) @ MAT INPUT A,C
30 MAT B=TRN(A)*A @ MAT D=TRN(A)*C @ MAT X=SYS(B,D) @ MAT DISP X;
```

-- Antonio

Thanks in advance

Re: Solving a linear system with physical meaning

Message #28 Posted by Valentin Albillo on 22 Jan 2008, 10:32 a.m.,
in response to message #27 by Antonio Maschio (Italy)

Hi, Antonio:

Antonio posteed:

"Well, if you don't mind, could you explain to me what's the purpose of the SYS() instruction?"

Yes, of course. The MAT SYS instruction solves M systems of N simultaneous linear equations in N unknowns.

For instance, MAT X=SYS(A,B) would solve the M systems with common coefficient matrix A and independent terms B, and would return the M solutions in matrix X.

In the simplest case of solving a single system, M=1 and both X and B are vectors or 1-column matrices.

Thanks for your interest and

Best regards from V

Re: Solving a linear system with physical meaning

Message #29 Posted by Stefan Vorkoetter on 22 Jan 2008, 12:03 p.m.,
in response to message #23 by Valentin Albillo

Thanks Valentin for clearing this up and finding the real problem.

I think what caused the belief that solving just the 4x4 system should be adequate was the fact that it **was** a 4x4 system. The fact that the system was square was arbitrary. After all, the number of calibration readings has nothing to do with the number of meters. It just so happened that 4 calibration readings were chosen for use with 4 meters, which resulted in a square matrix, which resulted in the belief that solving that 4x4 linear system should yield a meaningful answer.

In other words, because the choice of the number of readings made the system look like a nail (even though it wasn't), it created a strong desire to solve it using a hammer.

Stefan

Re: Solving a linear system with physical meaning

Message #30 Posted by Rodger Rosenbaum on 23 Jan 2008, 12:56 a.m.,
in response to message #29 by Stefan Vorkoetter

> Quote:
>
> Thanks Valentin for clearing this up and finding the real problem.
>
> I think what caused the belief that solving just the 4x4 system should be adequate was the fact that it **was** a 4x4 system.

If you re-read my posts, you will see that from the beginning I said that the mathematically exact solution was *not* adequate, and I asked for suggestions as to what might be a better solution, and how to find it. I was needling you guys to get out your calculators and play with some linear algebra.

> Quote:
>
> It just so happened that 4 calibration readings were chosen for use with 4 meters, which resulted in a square matrix,

It didn't just happen; I did it that way deliberately because that is just the situation which would lead people to the common belief that a standard linear system solution would be adequate.

> Quote:
>
> which resulted in the belief that solving that 4x4 linear system should yield a meaningful answer.

It may have resulted in such a belief for some people. I have learned to reject that belief and my whole purpose in this thread is to show why and when it should be rejected.

Re: Solving a linear system with physical meaning

Message #31 Posted by Stefan Vorkoetter on 22 Jan 2008, 12:05 p.m.,
in response to message #23 by Valentin Albillo

Actually, is it necessary to add all six additional equations. Isn't "nominally identical" transitive, so it would only be necessary to add three of the equations? For example, [1, -1, 0, 0], [0, 1, -1, 0], and [0, 0, 1, -1]?

Stefan

Re: Solving a linear system with physical meaning

Message #32 Posted by Valentin Albillo on 22 Jan 2008, 1:59 p.m.,
in response to message #31 by Stefan Vorkoetter

Hi, Stefan:

Stefan posted:

"Actually, is it necessary to add all six additional equations. Isn't "nominally identical" transitive, so it would only be necessary to add three of the equations? For example, [1, -1, 0, 0], [0, 1, -1, 0], and [0, 0, 1, -1]"

Well, true "identicity" is indeed transitive and just 3 equations would suffice. But here we're dealing with "nominal identicity", i.e., the meters are nominally assumed to be identical but in all probability they aren't, at least exactly (else all weights would be 1/4, period) so their degrees of identicity are also subject to some unknown variability which is to be taken into account as part of the optimization. That's why 6 extra equations are actually necessary.

Using your 3 equations instead, the weights are also close to 1/4, but distinct from the ones using six equations and obviously "less balanced", i.e., more unequal. Check weight #1 (.2557) and weight #4 (.2492) for instance, versus .2532 and .2518 respectively.

Thanks for your appreciation and

Best regards from V.

Re: Solving a linear system with physical meaning

Message #33 Posted by Stefan Vorkoetter on 22 Jan 2008, 3:54 p.m.,
in response to message #23 by Valentin Albillo

Just an added observation:

I just noticed that Valentin's computed weights don't add up to 1.0, but then realized that that is perfectly acceptable. Looking at the original input data, one column at a time, each meter individually tends to read low on average, so it's not surprising that the weight for each meter would be slightly more than 1/4.

Stefan

Re: Solving a linear system with physical meaning

Message #34 Posted by Karl Schneider on 23 Jan 2008, 12:10 a.m.,
in response to message #33 by Stefan Vorkoetter

Hi, Stefan --

You beat me to it! I saw that this morning, but didn't have time to prepare a post. Indeed, that is the reason for the elevated weights.

The sum of the ten underestimates is -0.56; the sum of the five overestimates is 0.22, and one measurement was exactly correct.

-- KS

Edited: 23 Jan 2008, 12:13 a.m.

Re: Solving a linear system with physical meaning

Message #35 Posted by Rodger Rosenbaum on 23 Jan 2008, 1:02 a.m.,
in response to message #34 by Karl Schneider

This isn't unexpected, and it isn't that the meters are reading low. I only used 16 random samples when I created the perturbed A matrix, and I wouldn't expect them to have a zero mean. Not with only 16. :-(

That's just like the real world; it takes a lot of samples for the mean to drift down to zero.

Re: Solving a linear system with physical meaning

Message #36 Posted by Rodger Rosenbaum on 23 Jan 2008, 12:41 a.m.,
in response to message #23 by Valentin Albillo

Quote:

There's nothing in your mathematical model that accounts for the empirical fact that you know for sure the meters to be "nominally identical", so given your readings corrupted with random errors, any set of weights is mathematically and logically as acceptable as any other.

I think you've attributed meaning to the "nominally identical" phrase that it doesn't have. Let me change my description of the meters. Suppose that the 4 (nearly perfect) meters have full scale readings of 5, 6, 7 and 8 amperes respectively. Now they are no longer indentical in any sense except that they read the same for

any current except for a small random error. Also suppose that their readings are unaffected by any scale factor error, or any other systematic error. Assume the only errors are additive noise. I created the A matrix in the original post by starting with an ideal matrix with identical columns of [1 2 3 4 } T (which could have come from 4 perfect, identical meters, or just as well could have come from 4 perfect, non-identical meters; I probably shouldn't have said "nominally identical". It seems to have engendered a complication the problem doesn't have) and adding random noise. It was the starting matrix of identical columns that I intended to imply by the phrase "nominally identical meters".

Now what you called a "known fact" no longer obtains. How does this change how you will determine the weights?

I don't agree that any set of weights is "mathematically and logically as acceptable as any other." Since the contaminating noise rendered the A matrix non-singular there is only one such solution. And, the "nominally identical" phrase didn't make it untrue. Surely you wouldn't say there was no mathematically acceptable set of weights when the meters were "nominally identical", but now that the meters are 5, 6, 7 and 8 amperes full scale, there is? Or do you say that there is still no mathematically acceptable solution even with the different meters?

I've reprinted Valentin's description with a literal variable replacing all the 1s in the last 6 rows:

Quote:

```
A(1,1)?     .95, 1.09, 1.01, 1.09
     A(2,1)?   1.94, 1.95, 2.02, 1.91
     A(3,1)?   2.92, 3,    2.93, 2.98
     A(4,1)?   4.01, 3.95, 3.97, 3.94
     A(5,1)?   k,  0,  0, -k          1st meter nominally identical to 4th
     A(6,1)?   k,  0, -k,  0          1st meter nominally identical to 3rd
     A(7,1)?   k, -k,  0,  0          ...
     A(8,1)?   0,  k,  0, -k          ...
     A(9,1)?   0,  k, -k,  0          ...
     A(10,1)?  0,  0,  k, -k          3rd meter nominally identical to 4th
```

Now let's evaluate this overdetermined system with various values for k:

```
 k        Solution
zero [ .882259887006 3.14485875706 -.483615819209 -2.54824858757 ]
.001 [ .857310152818 3.01809812277 -.442900447469 -2.43661963188 ]
.01  [ .406948019570 .711233605155 .289045001476  -.399771224958 ]
.1   [ .293670322429 .223226381774 .295373835235   .198164433485 ]
1    [ .253280216940 .252112563361 .253205061818   .251830755168 ]
10   [ .252613783902 .252602035207 .252613016547   .252599212812 ]
100  [ .252607078447 .252606960953 .252607070772   .252606932729 ]
1000 [ .252607011389 .252607010214 .252607011312   .252607009932 ]
10000 [ .252607010720 .252607010708 .252607010719   .252607010706 ]
```

Note that if we take the original A matrix:

```
[[ 0.95 1.09 1.01 1.09 ]
 [ 1.94 1.95 2.02 1.91 ]
 [ 2.92 3.00 2.93 2.98 ]
 [ 4.01 3.95 3.97 3.94 ]]
```

and post multiply it by:

```
[[ 1 ]
 [ 1 ]
 [ 1 ]
 [ 1 ]]
```

we get:

```
[[ 4.14 ]
 [ 7.82 ]
 [ 11.33 ]
 [ 15.87 ]]
```

Using this as an A matrix, and with our original B matrix, [1 2 3 4]T, we can solve as an underdetermined system. We get [[.252607010711]]T. If we multiply [[4.14 7.82 11.83 15.87]]T * [[252607010711] we get [1.04579302434 1.97538682376 2.98834093671 4.00887325998]T. This is the solution where all the weights are constrained to be equal (I gave this solution in another one of my posts), as we shall see.

If we divide [4.14 7.82 11.33 15.87]T by 4, we get [1.035 1.955 2.9575 3.9675]T. Now, create a new A matrix with 4 identical columns equal to the one we just calculated:

```
[[ 1.035  1.035  1.035  1.035 ]
 [ 1.955  1.955  1.955  1.955 ]
 [ 2.9575 2.9575 2.9575 2.9575 ]
 [ 3.9675 3.9675 3.9675 3.9675 ]]
```

If we try to solve a system with this A matrix using a standard linear solver, we'll get an error because this matrix is singular, and there are an infinite number of solutions. However if we create a solution vector (W column matrix) with the value we calculated a while back, namely [.252607010711 .252607010711 .252607010711 .252607010711]T, and use it to postmultiply this A matrix, we get [1.04579302434 1.97538682376 2.98834093671 4.00887325998]T, the same result as above.

This is the same result Valentin's procedure gets when we let the value of k become large.

Realizing that whenever we solve a linear system A*W=B, the data in the A matrix are constraints on the solution, we see that the additional constraints in Valentin's system when k = 1 are rather weak by comparing the value 1 to the other values in the first 4 rows. His constraints do tend to push the solution

toward a state of equality among the columns, but they don't push very hard. The parameter k must be quite large to get substantial equality.

I must say that I rather like Valentin's procedure. It provides a way to move continuously from the mathematically exact solution (with k=0) to the identical weights solution (with k=infinity as the limit), giving in-between solutions by selecting various values of k. There is at least one other method, by the way.

And there is no need to postulate "nominally identical" meters for the procedure to have merit. The procedure does its good work regardless of the nature of the meters. We need only know that the meter readings are from meters perfect in every way except for small random errors.

There may be desireable properties for some of those in-between solutions, such as less error magnification in exchange for greater residual norm. This is the very sort of thing I was trying to motivate people to explore, and it still hasn't been done.

Valentin gets a lot of credit for having actually done some calculations; now how about the rest of you guys?

Valentin said:

> Quote:
>
> ---
>
> which are the exact weights you found so unexpectedly (give or take a few ulps).
>
> ---

> Quote:
>
> ---
>
> which are the kind of weights you were expecting to begin with.
>
> ---

I don't know why so many people keep attributing this to me. I never said or implied any such thing. I *expected* the solution (using a standard linear solver) to be just what I got:

```
[[  .8822598871   ]
 [ 3.1448587566   ]
 [ -.4836158193   ]
 [ -2.5482485873 ]]
```

What I'm trying to get across in this thread is that most people would probably solve the system that way and think it was a good solution. I'm not one of those people. I've learned better and I'm trying to pass along what I've learned.

I explained why this isn't a good solution for this particular physical configuration. It has some undesireable properties, and there are other *approximate* solutions with desireable properties, and the error due to the fact that they are approximate is of the same order as the input data. I then asked how to find

them. So far, only Valentin has given a useful answer with numerical results showing how his procedure works.

Stefan said:

> Quote:
> _____
>
> Isn't "nominally identical" transitive, so it would only be necessary to add three of the equations? For example, [1, -1, 0, 0], [0, 1, -1, 0], and [0, 0, 1, -1]?
> _____

and Valentin pointed out that the result isn't quite the same. I would point out that if instead of using 1 in those 3 rows, use 10000 as I did in Valentin's system, and it becomes clear that as this parameter becomes very large, the solution tends to the same solution as Valentin's system when k is very large, which is the constrained equal weights solution.

Let's see ...

Message #37 Posted by [Valentin Albillo](#) on 23 Jan 2008, 7:53 a.m.,
in response to message #36 by Rodger Rosenbaum

Hi again, Rodger:

Rodger posted:

> *"I think you've attributed meaning to the "nominally identical" phrase that it doesn't have."*

> What part of *"nominally identical"* don't you understand ?

> *"Let me change my description of the meters."*

> Changing the stated conditions of a problem *a posteriori* (i.e. post-fact) is a big no-no in my book

> *"Suppose that the 4 (nearly perfect) meters have full scale readings of 5, 6, 7 and 8 amperes respectively."*

> The *"nearly perfect"* part is new as well. Originally they were *"4 nominally identical meters [...] These meters are just simple analog panel meters."*. I guess your new ranges for them mean that they are *"nominally identical"* no more as well ?.

"I probably shouldn't have said "nominally identical". It seems to have engendered a complication the problem doesn't have)"

Probably. And, by the way, it seems that it's just you the one seeing complications where actually there aren't any. You started with an inadequate, faulty model and got counterintuitive, faulty results. Correct your mathematical model and all complications and weird results disappear in thin air. That's all there's to it.

"It was the starting matrix of identical columns that I intended to imply by the phrase "nominally identical meters".

What you intended or not is simply not reflected in your wording of the problem's conditions. You said what you said and now pretend that everyone copes instead with what you *intended* to say,

"Now what you called a "known fact" no longer obtains. How does this change how you will determine the weights ?"

First of all, the unknown variables being computed are "weights", and the model I suggested does the rational thing, which is to assume your *"nominally identical"* condition essentially means "equal weight" in terms of weights, which is what we're computing here.

It doesn't matter whether the meters are the same model and color or have different scales, ranges, or sensitivities. It only matters that, for the purposes of your exposition, you deem they are *"nominally identical"* in the sense that their measuring accuracy and fiability are nominally equal, thus they can be considered to have the same weight as a working hypothesis.

That's what the improved model I suggested does. It explicitly specifies that every model should ideally have the same weight (i.e., measuring accuracy and fiability) as any other.

"I've reprinted Valentin's description with a literal variable replacing all the 1s in the last 6 rows:"

You should be aware that in most physical problems, there are a number of suitable models, not just the one, which if they are good enough, will simply get you similar results and the choice between one model or the other is left to either additional considerations (which can be reflected in the model in order to make it even more accurate) or else personal criteria.

For instance, in my suggested model I translated your *"nominally identical"* condition to "every meter nominally has the same weight as any other", and that did produce reasonable results.

But you could have also translated the constraint as "each meter is nominally identical to every other meter so their weights should reasonably be about 1/4", which would substitute my 6 extra equations with these four:

```
w1 = 1/4, w2 = 1/4, w3 = 1/4, w4= 1/4
```

and the least-squares solution would reflect this fact and also compute reasonable weights that would tend to fit the readings and tend to be 1/4 as closely as possible.

The selection of the correct model is thus not necessarily unique, and different correct models will give very similar results.

"This is the same result Valentin's procedure gets when we let the value of k become large [...] we see that the additional constraints in Valentin's system when k = 1 are rather weak by comparing the value 1 to the other values in the first 4 rows. His constraints do tend to push the solution toward a state of equality among the columns, but they don't push very hard. The parameter k must be quite large to get substantial equality.

Your modification of my suggested model to give it an additional degree of freedom accomplishes actually nothing as it's only a case of what I've already stated: there are diverse correct models, and in absence of more information, the choice is almost personal, all of them will give very similar, good results.

In this case, by inserting a constant k in the extra equations implementing the "equal weights" presuposed condition, you are just stating how much you care for fitting the readings versus having equal weights.

If you select k=0, you're telling the model: "I don't care for equal weights at all, please fit the readings and that's it" and you get the strange weights you didn't like.

if you select k=1000 or greater, you're telling the model: "I care very little for the readings, please make the weights equal" and that is exactly what you're getting, the weigths tend to be almost equal and the readings be damned.

"There may be desireable properties for some of those in-between solutions, such as less error magnification in exchange for greater residual norm. This is the very sort of thing I was trying to motivate people to explore, and it still hasn't been done."

In the absence of any suitable criteria for deciding how much you care for the readings and how much you care for equal weights, I simply selected the more or less balanced value of k = 1, which gives some importance to both data fitting and nominal weight equality.

Absolutely *nothing* in your initial conditions gives the slightest clue about your personal position in this matter, whether you'd swear over a pile of Bibles that the meters are nearly identical or whether you think that after taking the readings the least you can do is to pay some attention to them and attempt to fit them as well, just in case.

So, not being you, I can't tell which is which, and arbitrarily select k=1 as the most natural thing to do, which is neither better nor worse than k = 0.9 or k = 1.1, say.

You see, in the end it all depends now on your internal knowledge of the situation (how trustworthy are the meters, how identical, how trustworthy are the readings) and your personal choice for the free parameters, exclusively based in *that* subjective knowledge. No algorithm can take decisions for you.

"What I'm trying to get across in this thread is that most people would probably solve the system that way and think it was a good solution. I'm not one of those people. I've learned better and I'm trying to pass along what I've learned."

If that was your intention let me say that I think you didn't exactly succeed in conveying it to the readers, quite the opposite. There's a Spanish idiom for this but it doesn't translate well.

Best regards from V.

Edited: 23 Jan 2008, 8:05 a.m.

Re: Let's see ...

Message #38 Posted by [Rodger Rosenbaum](#) on 26 Jan 2008, 2:04 a.m.,
in response to message #37 by Valentin Albillo

A computer crash kept me from responding for a couple of days.

> Quote:
>
> ---
>
> Your modification of my suggested model to give it an additional degree of freedom accomplishes actually nothing as it's only a case of what I've already stated: there are diverse correct models, and in absence of more information, the choice is almost personal, all of them will give very similar, good results.
>
> ---

Are you suggesting that all of the various sets of weights obtained as the parameter k varies from 0 to inf all give similar results?

What would be the "results" which you would expect of the weights, to what use would you put them, that leads you to say that "all of them will give very similar, good results."

What mathematical test would you use to show that "all of them will give very similar, good results."?

Solving a linear system: Calculations

Message #39 Posted by Karl Schneider on 24 Jan 2008, 1:53 a.m.,
in response to message #36 by Rodger Rosenbaum

> Quote:
>
> Valentin gets a lot of credit for having actually done some calculations; now how about the rest of you guys?

Well, for what it's worth, I calculated the determinant and system solution of your original problem on an HP-42S, and got similar results within a few ULP's of those from your HP-48G* and from Valentin's HP-71B/Math ROM, but matching neither. It seems as though H-P was continually refining the linear-algebra code for the Saturn-based machines.

As far as I'm concerned, Valentin's solution handles the problem you posed to the furthest justifiable extent. If I had available simultaneous and nominally-identical measurements of a single quantity, I'd calibrate the devices and weight the measurements equally. If I wanted robustness from redundancy, I'd probably buddy-check the measurements against each other, throw out any bad one(s), and weight the good ones equally.

Regards,

-- KS

Edited: 24 Jan 2008, 3:11 a.m.

Re: Solving a linear system: Calculations

Message #40 Posted by Rodger Rosenbaum on 26 Jan 2008, 5:50 a.m.,
in response to message #39 by Karl Schneider

> Quote:
>
> As far as I'm concerned, Valentin's solution handles the problem you posed to the furthest justifiable extent.

The problem I gave was simple and symmetrical on purpose, and if the goal was just to solve that particular problem, no powerful mathematics is necessary at all. I gave the solution [.25 .25 .25 .25]T in my first post.

And as you say:

> Quote:
>
> If I had available simultaneous and nominally-identical measurements of a single quantity, I'd calibrate the devices and weight the measurements equally.

If the errors are small enough that the nature of the data {nonimally identical measurements) is apparent by visual inspection, the appropriate weighting can probably be determined without doing any matrix arithmetic at all.

But surely there must be some criterion for the "goodness" of a solution that is a little more mathematical than just how it looks (how close is a solution to the [.25 .25 .25 .25]T solution). Do you have any ideas?

But what if the system is not so symmetrical? Or what if it has a singular A matrix? Can we find an approximate solution?

I'll give a system here and start a new thread about it this weekend.

Given A:

```
[[ 4. 2. 3. 3. ]
 [ 5. 2. 6. 6. ]
 [ 9. 2. 5. 3. ]
 [ 4. 9. 4. 9. ]]
```

and B:

```
[[ 22.1743265837 ]
 [ 33.8228271732 ]
 [ 34.6819580525 ]
 [ 51.0426400707 ]]
```

Solve the system A*X=B

Re: Solving a linear system: Calculations

Message #41 Posted by Stefan Vorkoetter on 26 Jan 2008, 9:04 a.m.,
in response to message #40 by Rodger Rosenbaum

Quote:

```
[[ 4. 2. 3. 3. ]
 [ 5. 2. 6. 6. ]
 [ 9. 2. 5. 3. ]
 [ 4. 9. 4. 9. ]]
```

and B:

```
[[ 22.1743265837 ]
 [ 33.8228271732 ]
 [ 34.6819580525 ]
 [ 51.0426400707 ]]
```

Solve the system A*X=B

My Matrix Multi-tool program computes a determinant of -6.63e-10 and an X vector of [-8.1365, 17.4676, 26.5950, -20.0000].

However, the actual determinant is 0, so the system is singular and cannot be solved.

Stefan

Re: Solving a linear system: Calculations

Message #42 Posted by Rodger Rosenbaum on 26 Jan 2008, 9:03 p.m.,
in response to message #41 by Stefan Vorkoetter

Just because the determinant is 0 doesn't meant the system has no solution. In this case there are an infinite number of solutions.

If flag -22 is clear on the HP50, using the technique of putting B on level 2 of the stack, A on level 1 and pressing /, you get an error.

On the other hand, if flag -22 is set, then the HP50 gives:

```
[[ -1628.13646975 ]
 [ 2447.46761762 ]
 [ 4076.59499011 ]
 [ -3530.0000000 ]]
```

It's having the same problem as your program. Its internal LU decomposition is probably not getting a determinant of exactly zero.

Re: Solving a linear system: Calculations
Message #43 Posted by PeterP on 26 Jan 2008, 11:27 a.m.,
in response to message #40 by Rodger Rosenbaum

for what its worth Roger, thanks for posting an interesting problem, this was a good thread again (albeit it seems we still like to move to bickering in between :-)

Thanks again!

Cheers

P

Re: Solving a linear system: Calculations
Message #44 Posted by Karl Schneider on 26 Jan 2008, 3:35 p.m.,
in response to message #40 by Rodger Rosenbaum

Rodger --

> Quote:
>
> If the errors are small enough that the nature of the data {nonimally identical measurements) is apparent by visual inspection, the appropriate weighting can probably be determined without doing any matrix arithmetic at all.
>
> But surely there must be some criterion for the "goodness" of a solution that is a little more mathematical than just how it looks (how close is a solution to the [.25 .25 .25 .25]T solution). Do you have any ideas?

How about this: If each meter or transducer gives a reading or output result equal to the following:

The exact value + an uncorrectable calibration error + random noise of uniform distribution,

Then, one should be able to calculate the respective probabilities of each device providing the most accurate result for any given measurement. The weights could be set equal to those probabilities.

This is more of a statistical exercise than a linear-algebra problem, and I would say that a reference on metrology should be consulted.

> Quote:

Given A:

```
[[ 4. 2. 3. 3. ]
 [ 5. 2. 6. 6. ]
 [ 9. 2. 5. 3. ]
 [ 4. 9. 4. 9. ]]
```

and B:

```
[[ 22.1743265837 ]
 [ 33.8228271732 ]
 [ 34.6819580525 ]
 [ 51.0426400707 ]]
```

Solve the system A*X=B

The singular matrix A defines an <u>under-determined</u> system, so there are *infinitely many* solutions. The HP-48/49/50 can return the one with the lowest Frobenius (Euclidian) norm, as indicated on page 14-15 of the *HP-48G Series User's Guide*:

On my HP-49G, I got

[2.01952510743 2.23362533599 1.20500296581 2.00465552820]^T

as that solution for X. A*X produces each value of your constant matrix B to within one ULP.

Note how much lower in "pythagorean magnitude" this solution is, compared to Stefan's. The HP-48/49/50's capability of calculating exact determinants for integer-valued matrices allows it to identify a singular (under-determined) matrix, so that the optimization can be performed. The HP-15C, HP-42S -- and Stefan's program -- produce very small non-zero results for the determinant. In the case of the HP calculators specifically, roundoff errors from calculation of the LU decomposition are to blame.

-- KS

Edited: 26 Jan 2008, 6:04 p.m.

Re: Solving a linear system: Calculations
*Message #45 Posted by Rodger Rosenbaum on 26 Jan 2008, 9:42 p.m.,
in response to message #44 by Karl Schneider*

> Quote:
>
> ---
>
> How about this: If each meter or transducer gives a reading or output result equal to the following:
>
> The exact value + an uncorrectable calibration error + random noise of uniform distribution,
>
> Then, one should be able to calculate the respective probabilities of each device providing the most accurate result for any given measurement. The weights could be set equal to those probabilities.
>
> ---

This is a fine idea, but there is much more to be said for the simpler case of random noise only. I think it's well to know how do deal with that case effectively first, and then the further complications.

> Quote:
>
> ---
>
> The singular matrix A defines an under-determined system, so there are infinitely many solutions. The HP-48/49/50 can return the one with the lowest Frobenius (Euclidian) norm, as indicated on page 14-15 of the HP-48G Series User's Guide:
>
> ---

Even though HP calls this an under-determined system, I think that could be confusing. Since there are 4 equations and 4 unknowns, it looks to be an exactly determined system at first glance. Because 1 of the columns is a linear combination of the other 3, the system is technically under-determined, so it's not incorrect to call it that, but I think there is a better description.

The terminology used in the mathematical literature for this kind of system is "rank deficient" if the A matrix is not square, and "singular" if it is. I sometimes use the term "rank deficient" even for a square matrix, because that's what it is, and this term reminds us of what we must do to deal with it.

> Quote:
>
> ---
>
> A*X produces each value of your constant matrix B to within one ULP.
>
> ---

On my HP50, I get exact equality. Do you get 1 ULP differences for some of the values?

I assume you used the LSQ command to get this result. It is well to note that flag -54 must be clear to get the result you got. It's interesting to see what you get if you set that flag.

> Quote:
>
> ---

The HP-48/49/50's capability of calculating exact determinants for integer-valued matrices allows it to identify a singular (under-determined) matrix, so that the optimization can be performed.

The HP48G and descendants correctly calculate determinants of integer-valued matrices (where the elements are approximate numbers) by virtue of setting "tiny elements" to zero, according to the setting of flag -54. This typically doesn't need to be done for small integer matrices, and in fact for this one, the setting of flag -54 doesn't matter, even though the elements of the matrix are approximate, not exact, numbers. For exact numbers, an integer determinant will be returned for any matrix of only integer elements, regardless of the setting of flag -54, and I sometimes use this as a check.

This doesn't apply to the LSQ command, anyway.

LSQ's method of solution isn't described in either the HP48GX user's guide or the AUR. However, a third party book titled "Linear Algebra Investigations with the HP48GX, by Donald LaTorre, says that the LSQ function uses orthogonal decompositions, either the QR or the LQ decomposition. (LaTorre was involved in the design of the HP48G.) The determinant doesn't enter into it, but the setting of flag -54 does have an effect.

I'm going to start a new thread to continue this.

Teminology of matrices and linear systems

Message #46 Posted by [Karl Schneider](#) on 27 Jan 2008, 6:05 p.m.,
in response to message #45 by Rodger Rosenbaum

Rodger --

Upon further review of my linear-algebra textbook as well as the primary HP-48G reference, it seems that each of us was not always using strictly-orthodox terminology of matrices and linear systems in all cases. A few things ought to be cleared up.

You stated,

Quote:

... Since there are 4 equations and 4 unknowns, it looks to be an exactly determined system at first glance. Because 1 of the columns is a linear combination of the other 3, the system is technically under-determined, so it's not incorrect to call it that, but I think there is a better description.

The terminology used in the mathematical literature for this kind of system is "rank deficient" if the A matrix is not square, and "singular" if it is. I sometimes use the term "rank deficient" even for a square matrix, because that's what it is, and this term reminds us of what we must do to deal with it.

The term *"rank"* applies only to a <u>matrix</u>. The rank of a matrix is the dimension of its range. The rank cannot exceed the number of rows or the number of columns of a matrix, whichever is fewer. A matrix can be not square, and still be full-rank. "Rank-deficient" is equally applicable to square or non-square matrices. Of course, only a square matrix can be singular.

Over-determined, under-determined, and *exactly-determined* apply only to systems, e.g., A**x** = **b**. It also seems counterintuitive to me that an under-determined system can be made "exactly-determined" or even "over-determined" by simply padding it with non-independent rows, but... oh, well.

An *over-determined linear system* is defined in my text, as well as in pages 14-14 and 14-15 of the *HP-48G Series User's Guide*, simply as one which has more equations than independent variables. However, it is possible for an over-determined system to have no solution, one unique solution, or infinitely many solutions -- it depends on the rank of the system matrix, and whether the system is *consistent*. So, to reiterate the basic conclusions:

- If the system matrix is full-rank, but the (over-determined) system is consistent, then redundant information is present. The unique solution of the system can be obtained after simply deleting enough rows to make it square and exactly-determined.

- If the system matrix is full-rank, but the (over-determined) system is <u>not</u> consistent, then the least-squares "best fit" solution must be obtained.

I can't say with certainty what kind of solution(s) could be obtained with an over-determined and inconsistent system having a rank-deficient matrix.

BTW, flag -54 was indeed clear (the default setting) on my HP-49G when I obtained the least-squares "LSQ" solution to your exactly-determined system with the singular (thus, rank-deficient) matrix:

2.01952510743
2.23362533599
1.20500296581
2.00465552820

This solution, when multiplied by your matrix A, produced your vector B with an error of one ULP on two elements, and was exact on the other two.

I got a very different result from "LSQ" with flag -54 set:

673.501526294
-1004.98937644
-1677.5
1456.88232477

This solution, when multiplied by your matrix A, produced a residual from B of $10^{-10}*[263\ 368\ 275\ 493]^T$.

-- KS

Edited: 28 Jan 2008, 3:55 a.m.

Re: Teminology of matrices and linear systems

Message #47 Posted by Rodger Rosenbaum on 28 Jan 2008, 8:09 a.m.,
in response to message #46 by Karl Schneider

> Quote:
>
> > Quote:
> >
> > ... Since there are 4 equations and 4 unknowns, it looks to be an exactly determined system at first glance.
> > Because 1 of the columns is a linear combination of the other 3, the system is technically under-determined,
> > so it's not incorrect to call it that, but I think there is a better description.
> >
> > The terminology used in the mathematical literature for this kind of system is "rank deficient" if the A matrix is
> > not square, and "singular" if it is. I sometimes use the term "rank deficient" even for a square matrix, because
> > that's what it is, and this term reminds us of what we must do to deal with it.

You got me!

I was in a big hurry when I wrote that post and I left out the word "matrix" after the word "system" in the second sentence. I sometimes see the A matrix called the "design matrix" also.

Maybe I'll just call it a "square" matrix when there as many equations as variables.

On re-reading those two sentences, I realize I left out something. A rank deficient square system may have two natures depending on whether the rectangular matrix which is the A matrix augmented by the B matrix has a different rank than the A matrix alone. In one case it has an infinite number of solutions, and may be described as under-determined. In the other case, it has no solutions (is inconsistent) and may be called over-determined.

So a system with a square A matrix may be over-, exactly-, or under-determined in this sense.

> Quote:
> _____
>
> An overdetermined linear system is defined in my text, as well as in pages 14-14 and 14-15 of the HP-48G Series User's Guide, simply as one which has more equations than independent variables.
>
> _____

Actually, the User's Guide says, "These systems have more **linearly independent** (emphasis mine) equations than independent variables."

That "linearly independent" part is crucial. A system with more total equations than independent variables *could* have more or fewer "linearly independent" equations than independent variables.

> Quote:
> _____
>
> If the system matrix is full-rank, but the (overdetermined) system is consistent, then redundant information is present. The unique solution of the system can be obtained after simply deleting enough rows to make it square and exactly-determined.
>
> _____

I'm going to have to think about this one. Let's say you need to delete 3 rows to make the system square. Is the solution to the system with the deleted rows the same no matter which rows are deleted?

==

I get slightly different numerical results sometimes! I'm going to have to try this out on an HP48G and also the HP49G.

For example:

Quote:

This solution, when multiplied by your matrix A, produced a residual from B of 10^10*[263 368 275 493]T.

I get 10^-10*[223 368 235 753]T on the HP49G+.

Re: Teminology of matrices and linear systems

Message #48 Posted by Karl Schneider on 29 Jan 2008, 2:20 a.m.,
in response to message #47 by Rodger Rosenbaum

Quote:

Actually, the User's Guide says, "These systems have more **linearly independent** (emphasis mine) equations than independent variables."

That "linearly independent" part is crucial. A system with more total equations than independent variables could have more or fewer "linearly independent" equations than independent variables.

I'm still struggling to reconcile my textbook's *complete* statement with H-P's statement. Both might be correct, if properly interpreted (which I might not have done). There are different ways of stating things...

Quote:

If the system matrix is full-rank, but the (overdetermined) system is consistent, then redundant information is present. The unique solution of the system can be obtained after simply deleting enough rows to make it square and exactly-determined.

I'm going to have to think about this one. Let's say you need to delete 3 rows to make the system square. Is the solution to the system with the deleted rows the same no matter which rows are deleted?

Good point -- no, the square "reduced" matrix would need to remain full-rank, leaving a basis for the row and column space intact. In other words, only rows whose removal would not reduce the rank of the matrix should be deleted. For example, to leave two rows in the matrix that are merely scalar multiples of each other would reduce its rank.

-- KS

Edited: 29 Jan 2008, 2:25 a.m.

A slightly-belated "welcome back" (matrix solutions)

Message #49 Posted by Karl Schneider on 24 Jan 2008, 12:55 a.m.,
in response to message #23 by Valentin Albillo

Hi, Valentin --

A sincere "Welcome back!" from your 12-week hiatus, and thank you for the straightforward and clear analysis of the mathematical problem posed by Rodger. No obvious way to mathematically derive reasonable weightings was immediately apparent to me, but yours made perfect sense. This kind of insight is especially valuable.

> Quote:
>
> To that effect, I've just concocted the following 3-liner for the HP-71B:
>
> ```
> 10 DESTROY ALL @ OPTION BASE 1 @ INPUT "#E,U=";M,N
> 20 DIM A(M,N),C(M),B(N,N),D(N),X(N) @ MAT INPUT A,C
> 30 MAT B=TRN(A)*A @ MAT D=TRN(A)*C @ MAT X=SYS(B,D) @ MAT DISP X;
> ```
>
> which asks for the number of equations (E) and the number of unknowns (U), then lets you input the known data and proceeds to print out the computed values for the unknowns using least squares methodology.

Your elegant HP-71 program also reminded me of a theorem of linear algebra that needed to be "dusted off" after the passage of 12 or 17 years since I'd seen it. Perhaps other readers would benefit from an interpretation and explanation:

The program utilizes the theorem that, for an overdetermined and inconsistent linear system $A\mathbf{x} = \mathbf{b}$,

the solution $\mathbf{x}^*$ of $A^TA\mathbf{x}^* = A^T\mathbf{b}$ is the least-squares "best fit" solution. (Of course, if A is square and full-rank, then the solution will still be $\mathbf{x}$.)

On pages 14-14, 14-15, and 18-12 of the *HP-48G Series User's Guide*, we see that the HP-48G will automatically return optimized solutions for systems that are not exactly-determined:

- For overdetermined systems, the closest "solution" having the *least squares error* will be returned.

- For underdetermined systems, the solution having the lowest Euclidean (or Frobenius) "root-mean-square" norm -- namely, the *least squares magnitude* -- will be returned.

The HP-71 Math Pac Owner's Manual states on page 87 that non-square system matrices (and possibly the constant vector/matrix) must be padded accordingly with zeroes in order to make them square, so that "a solution" can be provided. (Of course, $A^T A$ is also square.) No information is given regarding whether these solutions are optimized.

So, with the "LSQ" (least squares) function, the HP-48/49/50 will automatically perform such optimizations, which must be done manually on an HP-15C, HP-42S, or (as Valentin did) on an HP-71B.

-- KS

Edited: 27 Jan 2008, 5:11 p.m. after one or more responses were posted

---

Re: A slightly-belated "welcome back" (matrix solutions)

Message #50 Posted by Valentin Albillo on 24 Jan 2008, 8:40 a.m.,
in response to message #49 by Karl Schneider

Hi, Karl:

Karl posted:

> *"A sincere "Welcome back!" from your 12-week hiatus, and thank you for the straightforward and clear analysis of the mathematical problem posed by Rodger. No obvious way to mathematically derive reasonable weightings was apparent to me, but yours made perfect sense. This kind of insight is especially valuable."*

> Thank you very much for both your fond welcome-back and your kind comments re my analysis, much appreciated. As for the "12-week hiatus", I've been dedicating my free time to some other hobbies o'mine having little to do with calcs or math and much to do with fine arts and ancient cultures (nihon), all in all extremely rewarding (if expensive).

> I did occasionally read this forum but was unlucky to find it infested by the usual suspects and their moronic ramblings and tantrums, and seriously lacking in interesting contents so I didn't stay for long. It saddens me to say this but frankly, IMHO, the magic is decaying at an accelerated rate and I foresee my participation diminishing accordingly. :-(

> *"Your elegant HP-71 program [...]"*

Thanks again but it's far from elegant, it's just a quik'n'dirty piece of code created on the fly to test my expanded model. As written, it uses three extra matrices just because I wanted the original data to remain unaltered, in order to check the fit. If this isn't necessary, you can do without the extra matrices, like this:

```
10 DESTROY ALL @ OPTION BASE 1 @ INPUT "#E,U=";M,N
20 DIM A(M,N),C(M) @ MAT INPUT A,C
30 MAT C=TRN(A)*C @ MAT A=TRN(A)*A @ MAT C=SYS(A,C) @ MAT DISP C;
```

which is shorter, slightly faster, and takes much less RAM, thus it's more "elegant".

Thanks for your interpretation and explanation, which I'm sure most readers will welcome, and

Best regards from V.

Re: A slightly-belated "welcome back" (matrix solutions)

Message #51 Posted by Thor Lansen on 26 Jan 2008, 11:50 p.m.,
in response to message #50 by Valentin Albillo

Welcome back my friend Valentin! in the words of Mark Twain, I am glad to hear the reports of your death had been greatly exaggerated.

Regards, Thor

Re: A slightly-belated "welcome back" (matrix solutions)

Message #52 Posted by PeterP on 26 Jan 2008, 11:26 a.m.,
in response to message #49 by Karl Schneider

Thanks Karl, I always learn lot from your posts!

Cheers

P

Re: Solving a linear system with physical meaning

Message #53 Posted by PeterP on 26 Jan 2008, 11:25 a.m.,
in response to message #23 by Valentin Albillo

simply beautiful, Valentin.

Cheers

Peter

[Return to Index | Top of Index]

GTO Go back to the main exhibit hall