From: Ralf Marbach <Ralf.Marbach@vtt.fi>
To: Jeremy Grata <JGrata@mail.bico.com>
Date: Wed, Feb 23, 2000 9:07 AM
Subject: Re: question

Jerry:

y^ty is a scalar. Of course it is calculable.

You are right that reference "error" does not reduce the magnitude of the b-vector (not systematically, anyways, that's why I separated the reference error out into unimportant Eq13).
"S" is the scaling between the true concentration in the sample (in your case, skin) and the reference concentrations (in your case, blood). S has to be in the formula for the b-vector. If it was not, then your predictions would all be flat (since there is at least 3x less mg/dL glucose-in-skin than glucose-in-blood).

Hey, would you like to try the "modern" version of calibration on your non-invasive data? I have never tried it and it will be several days work, but just imagine, if you scaled the glucose-in-water spectrum, at each wavelength according to your best guess as to what the pathlength at that wavelength thru glucose-bearing tissue is, and if then the b-vectors modern and traditional way came out the same, you'd have plenty of reason to party. (With ground hog ground meat.) Just imagine the effect on the FDA (of the "physical" modeling, I mean, not the ground hogs.)

Best

CC: IND__PA.smtp("ralf.marbach@ele.vtt.fi")

From: Ralf Marbach <Ralf.Marbach@vtt.fi>
To: Jeremy Grata <JGrata@mail.bico.com>
Date: Thu, Feb 24, 2000 3:33 PM
Subject: Re: more help

Hi Jerry:

Here my comments:

> I'm trying to get my mathematician to give me a list of items that he disagrees
> with in your paper ... this guy is tough!

A whole list? Since when do mathematicians have to agree or disagree on anything, anyway? It is either correct or wrong. Did he find any mistakes?

> We're in the process of determining how to do SIC (using LS or Orthogonal
> regression). I'm inclined to believe that when calculating the self prediction
> slope that orthogonal regression is the best choice for the following reason:

There is no "best" slope, just what you like best.

> PLS sees Xn noise and S noise and pulls the b-vector down.

Yes right. (I call S a scaling factor, not noise.)

> When calculating the self pred slope, the x-axis noise further pulls the slope
> down. This is an artificial decrease in slope and should not be accounted for
> when slope correcting the b-vector. In that way, independent predictions have the
> correct sensitivity (true slope) but will have a slope less than one due to x-axis
> errorif calculated by LS regression and a slope of one if calculated by
> orthogonalregression (this is something we have seen).

The physicist's way to define "best" slope. I like it.

> In the limit of no x-axis (reference value) error, the self-pred slope should be
> the same for LS and orth. To test this we tried RBA calibrations which in
> theory should take care of x-axis error. The problem is that the orthogonal
> regression sees a slope greater than one for self-pred RBA. This suggests that
> RBA has some non-linear component increasing the b-vector or that orthogonal
> regression doesn't take care of x-axis noise properly (the latter is tough to
> accept since orhtognal regression is a sound regression tecnique).

Why should RBA "in theory" reduce reference noise? Assume reference noise was 100% dominated by lead/lag; then the reference noise increases linearly with distance from the reference mean value, and RBA should not do anything to diminish it, 100% ineffective.

> I'm still investigating the effects of LS and orthogonal on the profile data ...
> this may shed some more light.
>
> any comments?

If your goal is to distinguish spectral noise vs. reference noise in your data, do the following: determine the mean(SNR) of the type-1 and type-2 patients. The mean signal ratio is about 1.5 if I remember correctly, 75 mg/dL RMS over 50 RMS. Assume (a) SNRx(typ1)=(1.5)xSNRx(type2) but (b) SNRy(typ1)=SNRy(type2) , because of the lead/lag issue above. Insert (a) and (b) into the equation(18) and solve the two equations for type1 and 2 for the two unknowns SNRx and SNRy. I bet $10 the SNRx of type-1's is better than 2.5 !!

Ralf

From: "Marbach, Ralf" <marbach@cytometrics.com>
To: 'Jeremy Grata' <JGrata@mail.bico.com>
Date: Mon, Mar 8, 1999 11:31 AM
Subject: RE: Questions

hey Jerry,

keep vibrating. Here come answers:

1) Disregardig the HemoCue error (ie. no noise on the y-axis) and blaming it all on the spectral noise, SNR is then only a function of how much (1) RMS glucose signal (mg/dL RMS) you got on a patient, (2) glucose absorbance (AU /(mg/dL)), and (3) spectral noise (AU rms) in the glucose direction. The formulas are:

slope = SNR^2 / (1 + SNR^2)

corr coef = sqrt(slope)

scatter/signal = SNR / (1 + SNR^2)

Now, keep in mind that, yes, the Hemocue is accurate, but it is also measuring the wrong thing and it is therefore inaccurate w.r.t. to ISF glucose. Formulas above, which are on your plot, do not show that. I derived the slope for the case including y-noise and it is:

slope = ((yhcT*yt)/(ytT*ytT)) x SNR^2 / (1 + SNR^2)

where yhc and yt are the mean-centered vectors [mg/dL] for blood and tissue, respectively. In summa, the slope is determined by both the spectral noise (SNR) and the correlation blood-tissue, and that's why slopes shake out to be the same between type I and II's (you expect the I's to have higher slope because of more signal (mg/dL rms) but they don't).

2) b-vector sensitivity comes from the b-vector alone, you do not need dark data. Multiply b-vector with constant offset (1 mAU flat) to get offset sensitivity (mg/dL / mAU); take sqrt(sum(b.^2))/1e6 to get random noise sensitivity (md/dL / microAU); and multiply by wavelength-shifted version of mean skin absorbance spectrum to get wavelength shift sensitivity (mg/dL / nm). There is an M-file on my old PC and on grunt (bsen.m) that does that for you.

Say hi to everybody I know, and I hope you like snow,

Ralf

Ralf Marbach, Ph.D.
Cytometrics, Inc. marbach@cytometrics.com
615 Chestnut St., 17th Floor Voice 215.574.7002
Philadelphia, PA 19106 Fax 215.574.7099

-----Original Message-----
From: Jeremy Grata [mailto:JGrata@mail.bico.com]
Sent: Monday, March 08, 1999 8:51 AM
To: Marbach, Ralf
Subject: Questions

Hey man ... what's up!? Keepin busy?

It's busy around here. Some times feels like my brain is just vibrating.

I have 2 questions for you:

1. Where can I find the equations for the Slope and Scatter dependence on SNR plot that we had hanging on the wall?

2. What was the equation for determining the machine sensitivity in (mg/dL)/(uAU) given the b-vector and the dark data?

Well ... gotta run ... Monday morning meeting time (I hate those meetings!).

later

jer

From: "Marbach, Ralf" <marbach@cytometrics.com>
To: 'Jeremy Grata' <JGrata@mail.bico.com>
Date: Fri, Mar 12, 1999 1:24 PM
Subject: RE: another

Jerry,

I am not sure I understand your question. However, being a city person now, I will still give you an answer. (Screw communications!)

Glucose peak absorbance @ 1.6 micron and @ 1 mm pathlength is about 1 microAU / (mg/dL). In-vivo this value is washed down (roughly 3x I remember) because 1 dL of skin does not contain as much glucose as does 1 dL of blood.

Now, I do not recall Mark's calculation to be erroneous. The discrepancy probably comes from the fact that his numbers are Euklidean length number, i.e., they depend on over what spectral range / resolution the glucose spectrum is measured.

So, at 100 mg/dL, 3x washdown, 0.5 mm path, the peak absorbance @ 1.6 micron is 17 microAU. If you rms-up the total Euklidean length and include neighboring wavelenghts, they do not cosiderably contribute compared to the peak @ 1.6 micron. So, Mark's number sounds correct to me. (Sorry).

Regards,

Ralf

PS --- Say hi to Mike. Hope to talk to him soon.

Ralf Marbach, Ph.D.
Cytometrics, Inc.
615 Chestnut St., 17th Floor
Philadelphia, PA 19106
marbach@cytometrics.com
Voice 215.574.7002
Fax 215.574.7099

-----Original Message-----
From: Jeremy Grata [mailto:JGrata@mail.bico.com]
Sent: Friday, March 12, 1999 11:16 AM
To: Marbach, Ralf
Subject: another

I was reading some of Mark D's papers. He talks about the glucose100 spectrum, but assigns a value of 16uAU to it. He quotes Ed H's glucose and water tests as verification. My glucose and milk test and the PWS's on the patient data indicate that glucose100 is at least 160uAU. Is it safe to assume that Mark D's calculation is just an error?

jer

CC: "'rmarbach@bellatlantic.net'" <rmarbach@bellatlant...

From: "Marbach, Ralf" <marbach@cytometrics.com>
To: 'Jeremy Grata' <JGrata@mail.bico.com>
Date: Tue, Mar 16, 1999 10:28 AM
Subject: RE: rba

Jerry,

we got a little snow, not to mention.

Huh, RBA is a long time ago. I do not remember details but it goes s.th. like the lottery example below.

(German lottery: 6 out of 49) If you want to pick 6 numbers randomly out of the numbers [1:49] without(!) replacement, meaning you can't pick "7" twice just because you like it so much, then you got (49!)/(6! * 43!) combinations.
This formula does not consider the order of you picking, meaning "50 3 71", "3 71 50" etc. etc. only count as one combination. (Imaging otherwise: you could claim you hit the jackpot 6!=720 times at once and want all the money NOW.)

I used the above formula to figure things our with RBA.

Cheers,

Ralf

PS -- still waiting to hear from Mike.

Ralf Marbach, Ph.D.	
Cytometrics, Inc.	marbach@cytometrics.com
615 Chestnut St., 17th Floor	Voice 215.574.7002
Philadelphia, PA 19106	Fax 215.574.7099

-----Original Message-----
From: Jeremy Grata [mailto:JGrata@mail.bico.com]
Sent: Tuesday, March 16, 1999 9:40 AM
To: Marbach, Ralf
Subject: rba

There was a combinatorial calculation that we did to determine how many ++- combinations we could have (in rba) without over sampling. How did that calculation go?

did you guys get any snow this weekend?

From: Ralf Marbach <Ralf.Marbach@vtt.fi>
To: Jeremy Grata <JGrata@mail.bico.com>
Date: 4/17/00 11:24PM
Subject: Re: what the...

Jerry,

I looked thru your email. You are right, the SNR_i is independent of the b-vector, but that's allright. In fact, it's exactly what it should be.

(1) SNR is equivalent to correlation, i.e., it is a thing that exists in, say, 300-dimensional time-space.

(2) b-vector is a thing that exists in, say, 64-dimensional spectral space. When dot-multiplied into 300 spectra, it produces a time-signal which, in turn, has SNR.

(3) What you are doing is to measure the SNR of the individual u_i time-vectors. You don't need a b-vector to do that, you know what direction in time-space you are looking at, period.

Mathematically, and using your notation, you're measuring correlation between g and u_i. You can replace u_i with y_i=u_i * (s_ii *v_iT*b) using any b_vector you want, does not make a difference.

In the past, we looked at the u_i's with high SNR's and then made the jump to spectral space by looking at their v_i brothers, because we wanted to check those v_i's to look like the glucose absorbance. This makes sense because only those v_i's that have u_i's with high SNR's, will not be weighted down by the (1+SNR^2)-term and can make a contribution to the real b-vector when it is computed. But, for what you are doing, nobody cares about how the b-vector will look like.

Or in other words, the tough part is to find SNR in your data, computing the b-vector that goes along with it is easy.

Hope that helps,

Ralf

Jeremy Grata wrote:

> looked at the math, it gets you to the same equation. This is good since the
> results are the same (both algorithms give identical results). Do you have any
> data you can try this out on yourself? I attached the code I'm using (a is cal
> spectra, y is actual concentration values, b is bvector of interest).
>
> jer
>
> >>> Ralf Marbach <Ralf.Marbach@vtt.fi> 04/02/00 05:39PM >>>
> Jerry,
>
> Somewhere in your math obviously is a glitch that makes the b-vector cancel
> out.
> Your definition of SNR is confusing. Try computing the correlation coefficient
> of

> the "total response for rank i" and then traslate correlation coeff into SNR
> rank-byrank. That will work.
>
> Ralf
>
> Jeremy Grata wrote:
>
> > I don't know why I didn't see this problem before. Maybe my calculations are
> in
> > error, but check these and let me know. these equs come for the SNR
> > calculations we did while you were here:
> >
> > I define c=calibration spectra and g their corresponding finger pricks (both
> g
> > and c mean centered, g is normalized).
> >
> > u*s*v'=c (from svd)
> >
> > then I calculate the following:
> >
> > y=u_i * s_ii *v_i *b (the total glucose response for rank i)
> >
> > ys_i = g(g'*y_i) (the glucose signal for rank i)
> >
> > yn_i = y_i - ys_i (the glucose noise for rank i)
> >
> > and the snr for rank i it
> >
> > std(ys_i)/std(yn_i)
> >
> > but yn_i is actually
> >
> > y_i - g(g'*y_i) or (I - g*g') * y_i
> >
> > so the SNR for rank i is
> >
> > std (y_i) / std((I - g*g')*y_i)
> > this turns out to be independent of the bvector.
> >
> > I tested this experimentally and found that calculating snr vs rank this way
> was
> > independent of whatever b-vector I used.
> >
> > any suggestions?
>

```
> ------------------ code.txt follows ------------------
> y=(y-mean(y))/norm((y-mean(y)));
> [u,s,v]=svd((a-ones(size(a,1),1)*mean(a))/sqrt(size(a,1)),0);
> for i=1:size(a,2);ya(:,i)=u(:,i)*s(i,i)*sqrt(size(a,1))*v(:,i)'*b;end
> for i=1:size(a,2);yasignal(:,i)=y*(y'*ya(:,i));end;
> yanoise=ya-yasignal;
>
> snr=std(yasignal)./std(yanoise);
> snrsq=snr.^2;
> snr=[snr;snrsq];
>
```

```
> %next 5 lines compute snr via corrcoeff.
>
> x=corrcoef([y,ya]);
> rsq=x(1,2:size(a,2)+1).^2;
> snrsq=rsq./(1-rsq);
> snr=sqrt(snrsq);
> snr=[snr;snrsq];
```