

The modeling success and the solver success

Every time we run a data modeling program we have access to two publishable numbers $1 - |\mathbf{r}|/|\mathbf{d}|$ and $1 - |\mathbf{F}'\mathbf{r}|/|\mathbf{F}'\mathbf{d}|$. The first says how well the model fits the data. The second says how well we did the job of finding out.

Define the residual $\mathbf{r} = \mathbf{F}\mathbf{m} - \mathbf{d}$ and the “size” of any vector, such as the data vector, as $|\mathbf{d}| = \sqrt{\mathbf{d} \cdot \mathbf{d}}$. The number $1 - |\mathbf{r}|/|\mathbf{d}|$, will be called the “modeling success at fitting data.” (When the data fitting includes a residual weighting function, it should be incorporated in \mathbf{F} and \mathbf{d} .)

While the modeling success is of interest to everyone, the second number $1 - |\mathbf{F}'\mathbf{r}|/|\mathbf{F}'\mathbf{d}|$, to be called the “solver success at achieving goals,” is more of interest to us, the analysts. It tells us to what extent our program has achieved our stated goals of data fitting (and regularization). Experience seeing this number may give us guidance to where opportunities have been missed and where more work might be worthwhile. In problems of low dimensionality we can normally drive the solver success to unity. Conjugate-gradient theory says given infinite precision arithmetic, iteration should converge to the exact solution in a number of iterations equal to the number of unknowns. In reality, however, in Geophysics we often find ourselves iterating on problems far too large to run to completion. The solver success number tells us how well we are doing.

There are two interpretations of $\mathbf{F}'\mathbf{r}$. The most basic is that it is the residual \mathbf{r} , originally in data space, transformed to model space by \mathbf{F}' . The alternate interpretation of $\mathbf{F}'\mathbf{r}$ is that it is the gradient of the penalty function, namely $d(\mathbf{r} \cdot \mathbf{r})/d\mathbf{m}$. This gradient vanishes as the best fit is found. The statement that this gradient vanishes is the “normal equations”

$$\mathbf{0} = \mathbf{F}'\mathbf{r} = \mathbf{F}'(\mathbf{F}\mathbf{m} - \mathbf{d}) = (\mathbf{F}'\mathbf{F})\mathbf{m} - \mathbf{F}'\mathbf{d} \quad (1)$$

The gradient $\mathbf{F}'\mathbf{r}$ is also called $\Delta\mathbf{m}$ in many solver programs because of its role in building an update for the model \mathbf{m} . The progress of a solver can be observed by a plot of $|\Delta\mathbf{m}| = \sqrt{\Delta\mathbf{m} \cdot \Delta\mathbf{m}}$ versus iteration. Ideally it converges to zero. Unfortunately the amplitude axis of this plot scales with the units of model space. We should nondimensionalize it. Then we could make a simple statements like, “We iterated 90% of the way to the solution.”

Starting from a zero model $\mathbf{m} = \mathbf{0}$ the first residual is $\mathbf{r} = -\mathbf{d}$. Its size in model space is $|\mathbf{F}'\mathbf{d}|$ (which also happens to be our first estimated model). After many iterations the residual in model space is $|\mathbf{F}'\mathbf{r}|$. The ratio $|\mathbf{F}'\mathbf{r}|/|\mathbf{F}'\mathbf{d}|$ progresses towards zero as we progress. The degree of success our solver is having is $1 - |\mathbf{F}'\mathbf{r}|/|\mathbf{F}'\mathbf{d}|$. We could say, “Starting from our first model, we iterated until we had gone $100 \times (1 - |\mathbf{F}'\mathbf{r}|/|\mathbf{F}'\mathbf{d}|)\%$ of the way to the ultimate model \mathbf{m} .” More simply, $1 - |\mathbf{F}'\mathbf{r}|/|\mathbf{F}'\mathbf{d}|$ is the “solver success.”

The “modeling success at fitting data” is the number $1 - |\mathbf{r}|/|\mathbf{d}|$. The “solver success at achieving goals” is the number $1 - |\mathbf{F}'\mathbf{r}|/|\mathbf{F}'\mathbf{d}|$.