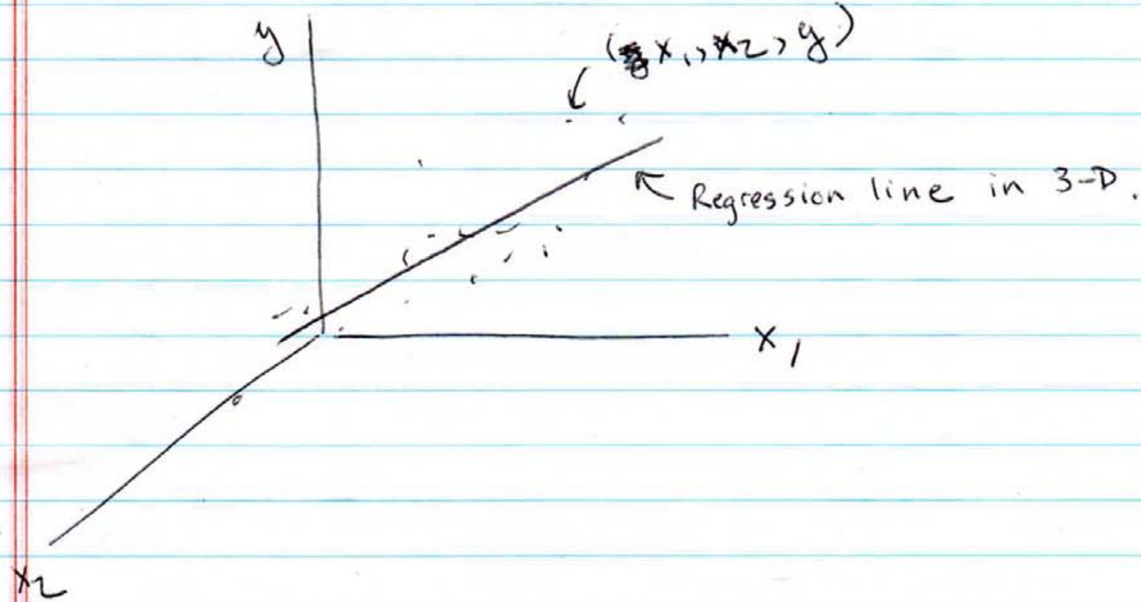
Multiple Regression

Idea: Generalize the derivation of the linear regression coefficient to multiple predictors.

$$\hat{y} = a_0 + a_1 x_i + a_2 x_2 + \dots a_N x_N$$

Fit now is in a multiple phase space.



If x_1 and x_2 are orthogonal predictors (i.e. they are not related) they give independent information regarding y.

Recall the expressions we got for the case of one predictor from our discussion of regression before

By least-squares error minimization, get:

$$\bar{y} = a_0 + a_1 \bar{x} \Rightarrow \text{Minimization wrt } a_0$$

$$\overline{xy} = a_0 \bar{x} + a_1 \overline{x^2} \Rightarrow \text{Minimization wrt } a_1$$

Now generalized to multiple predictors:

$$\bar{y} = a_0 + a_1 \bar{x}_1 + a_2 \bar{x}_2 + \dots + a_n \bar{x}_n \rightarrow \text{with respt. to } a_0.$$

Get multiple equations for minimization wrt a_1

For first predictor (x_1)

$$\overline{x_1 y} = a_0 \bar{x}_1 + a_1 \overline{x_1^2} + a_2 \overline{x_2 x_1} + \dots + a_n \overline{x_n x_1}$$

For n^{th} predictor (x_n)

$$\overline{x_n y} = a_0 \bar{x}_n + a_1 \overline{x_1 x_n} + a_2 \overline{x_2 x_n} + \dots a_n \overline{x_n^2}$$

If we assume normalized variables (i.e. mean = 0 and std. dev. = 1), then the equation reduces to:

$$\overline{(x_i x_j)}\, a_i = \overline{x_i y}^{\,*}$$

Covariance of predictors (2-D) MATRIX

Coefficients (1-D) VECTOR

Covariance of predictors and predictands (1-D) VECTOR

In matrix notation:

$$\begin{array}{cc} & j \rightarrow \\ i \downarrow & \begin{bmatrix} \overline{x_1^2} & \overline{x_1 x_2} & \overline{x_1 x_3} & \cdots & \overline{x_1 x_n} \\ \overline{x_2 x_1} & \overline{x_2^2} & \overline{x_2 x_3} & \cdots & \overline{x_2 x_n} \\ \overline{x_3 x_1} & \overline{x_3 x_2} & \overline{x_3^2} & \cdots & \overline{x_3 x_n} \\ \vdots & \vdots & \vdots & & \vdots \\ \overline{x_n x_1} & \overline{x_n x_2} & \overline{x_n x_3} & \cdots & \overline{x_n^2} \end{bmatrix} \end{array} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} \overline{x_1 y} \\ \overline{x_2 y} \\ \overline{x_3 y} \\ \vdots \\ \overline{x_n y} \end{bmatrix}$$

LHS: Covariance matrix of predictors

RHS: Covariance vector of predictors and predictands.

* If x's are standardized, LHS becomes the correlation matrix.

If the predictors are orthogonal, then the cross correlation terms (i.e. $\overline{x_1 x_2}$, etc.) are $= 0$.

Then the correlation coefficients are:

$$a_1 = \frac{\overline{x_1 y}}{\overline{x_1^2}} \;; \quad a_2 = \frac{\overline{x_2 y}}{\overline{x_2^2}} \quad \text{etc} \ldots$$

Otherwise must invert the covariance matrix to solve..

<u>Case of two predictors</u>

$$\hat{y} = a_0 + a_1 x_1 + a_2 x_2 \qquad \text{Regression equation}$$

If the variance terms $= 1$ (i.e. $\overline{x_1^2} = 1$ and $\overline{x_2^2} = 1$), then a_1 and a_2 are:

$$a_1 = \frac{\overline{x_1 y} - (\overline{x_2 y})(\overline{x_1 x_2})}{1 - (\overline{x_1 x_2})^2}$$

$$a_2 = \frac{\overline{x_2 y} - (\overline{x_1 y})(\overline{x_1 x_2})}{1 - (\overline{x_1 x_2})^2}$$

Can get this using matrix multiplication..
(possible HW problem)

If the variables (predictors + predictands) have been normalized; can express this in terms of correlation coefficients (r)

$$a_1 = \frac{r_{1y} - r_{12} r_{2y}}{1 - r_{12}^2}$$

$$a_2 = \frac{r_{2y} - r_{12} r_{1y}}{1 - r_{12}^2}$$

Calculate the explained variance and unexplained variance, then, in a similar way as with one predictor.

Additional notes

- Multiple Linear regression can be used as a data filter - and often is.

- Same caveats as with linear regression / correlation mentioned earlier.

- Can't just keep adding potential predictors blindly without evaluating how much they improve exp. variance.

Introduction to Part II of Course

As we've seen already, a fundamental skill we must have in doing geophysical analysis is the ability to analyze data in space.

Idea in this part of the course is to develop ~~■~~ methodologies for looking at statistically significant and physically interesting ~~■~~ spatial patterns.

Roadmap of this part:
- Review of matrix methods (linear algebra)
- EOF / PC analysis: sig. spatial patterns in data
- Discussion of other methods which relate to EOF analysis: SVD analysis, Canonical correlation analysis, etc.

→ Goal: Get to point where comfortable doing EOF analysis and can easily understand the other methods and the appropriate research applications.