Positive Definite and Semidefinite Matrices

The topic of this lecture

- All $\lambda_i > 0$
- Energy $x^TSx > 0$ (all $x \neq 0$)
- $S = A^TA$ (independent columns in A)
- All leading determinants > 0
- All pivots in elimination > 0

They all come together. Each one gives a test for positive definite matrices.

"**leading determinants**" means that we take the one by one determinant, it would have to be positive.

Deep learning, including neural nets, machine learning, the big computation in which we need to minimize the **energy** x^TSx

How to find a minimum point in

$$f(x, y) = x^TSx + x^Tb$$

We are trying to solve a big nonlinear system.

First we take a point, hopefully down the bowl. Second, we would compute the derivative with respect to x and y i.e. $\frac{df}{dx}$ and $\frac{df}{dy}$ or ∇f. Then we would do a part of **gradient descent** -- **line search** to find a second point and stop there. Recalculate the gradient, finding the steepest way down from that point, following it until it turns up or approximately, then you are at a new point. the whole process is called gradient descent -- the big algorithm of deep learning of neural nets, of machine learning, of optimization.

Notice that we didn't compute second derivatives. If we computed second derivatives, we could have a fancier formula that could account for the curve. But to compute second derivatives when you have got hundreds and thousands of variables is not a lot of fun. So most effectively, machine learning is limited to first derivatives -- the gradient. So that is the general idea. But there are lots and lots of decisions and it doesn't always work well. Let's see the what the trouble is.

It turns out, if you are going down a " narrow valley ". You take very small steps, just staggering back and forth across this and getting slowing, but too slowly, toward the bottom. So that's why things have got to be improved. If you have a small eigenvalue and a very large eigenvalue, those tell you the shape of the curved surface.

Go back to the original topic, let's think about a question -- Matrices S, T are positive definite matrices, what about $S + T$, is that matrix positive definite?

We look at the five test to find which one will be good to use. Let's look at the energy $x^T(S + T)x$, is that a positive number or not, for every x ?

Just separate those into two pieces

$$x^T(S + T)x = x^T(S)x + x^T(T)x$$

And both of those are positive, so the answer is yes.

What about S^{-1}, is that positive definite? the answer is yes because of the eigenvalue test.

What about Q^TSQ, Q is an orthogonal matrix. Notice

$$Q^TSQ = Q^{-1}SQ$$

and it is similar to S, so using the eigenvalue test, the answer is yes. actually we can also use energy to test it.

Let's introduce the semidefinite.