Statistics from Linear Algebra	**Math 21b, O. Knill**

RANDOM VARIABLES. In probability theory, functions are called **random variables**. They form a linear space. $E[f] = \int f\ dx$ is called the **expectation** of a random variable f, $Var[f] = E[(f - E[f])^2]$ is called the **variance** of f, $Cov[f,g] = E[(f - E[f])(g - E[g])]$ is the **covariance** of f and g.
One can think of $Cov[f,g]$ as the inner product of $f - E[f], g - E[g]$.

LINEAR REGRESSION. Given data points $(x_1, y_1), ..., (x_n, y_n)$ we can find the line $y = ax + b$ which best fits these points. We had the formula $\begin{bmatrix} a \\ b \end{bmatrix} = (A^T A)^{-1} A^T \vec{b}$, where $\vec{b} = \begin{bmatrix} y_1 \\ ... \\ y_n \end{bmatrix}$ and $A = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ ... & ... \\ x_n & 1 \end{bmatrix}$. Now
$A^T A =$