**BIRLA INSTITUTE OF TECHNOLOGY & SCIENCE, PILANI**

**BITS C464 – MACHINE LEARNING**

**I Semester 2014-2015**

**WORKSHEET #2**

**Linear Regression using Linear Basis Functions**

**OBJECTIVE:-**

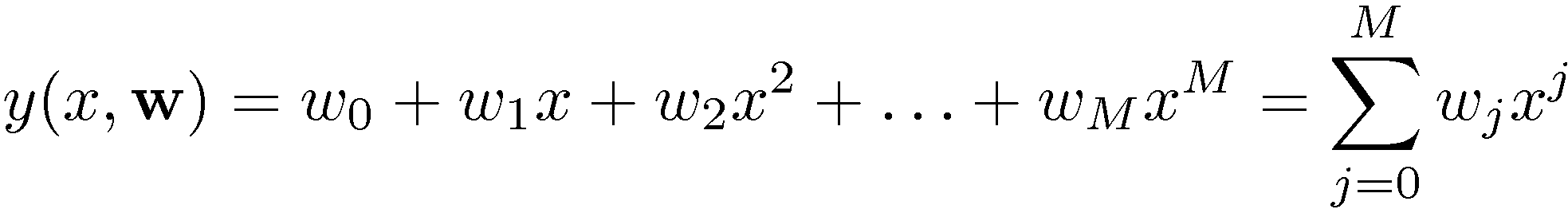
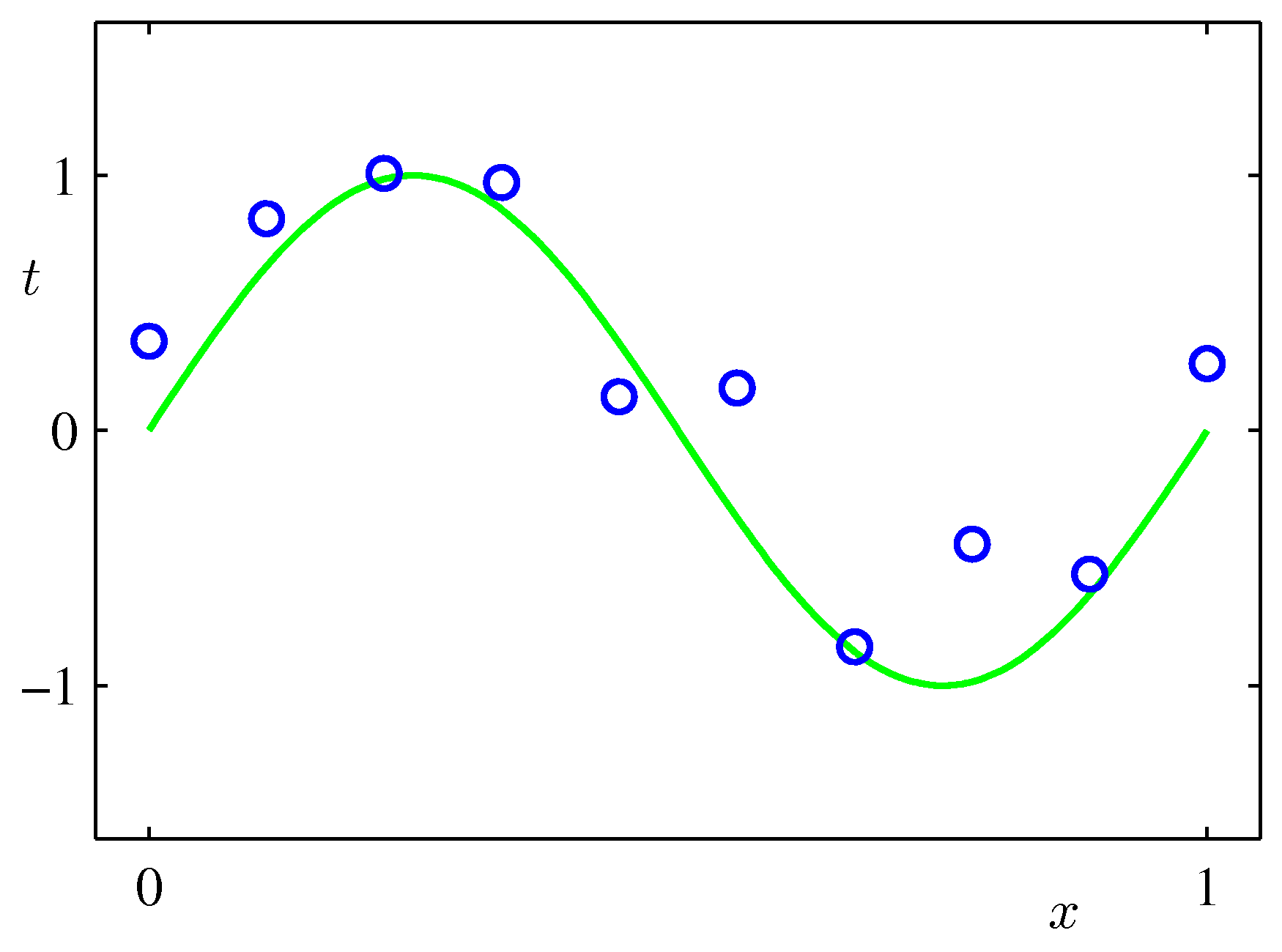
* Linear Basis function Models
* Different Basis Function Models as Polynomial, Gaussian, Sigmoidal
* Analysis of different basis function on the given data set(using Least Square Error)

**Linear Basis Functions**

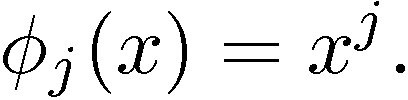
A **basis function** is an element of a particular basis for a function space. Every continuous function in the function space can be represented as a linear combination of basis functions, just as every vector in a vector space can be represented as a linear combination of basis vectors.

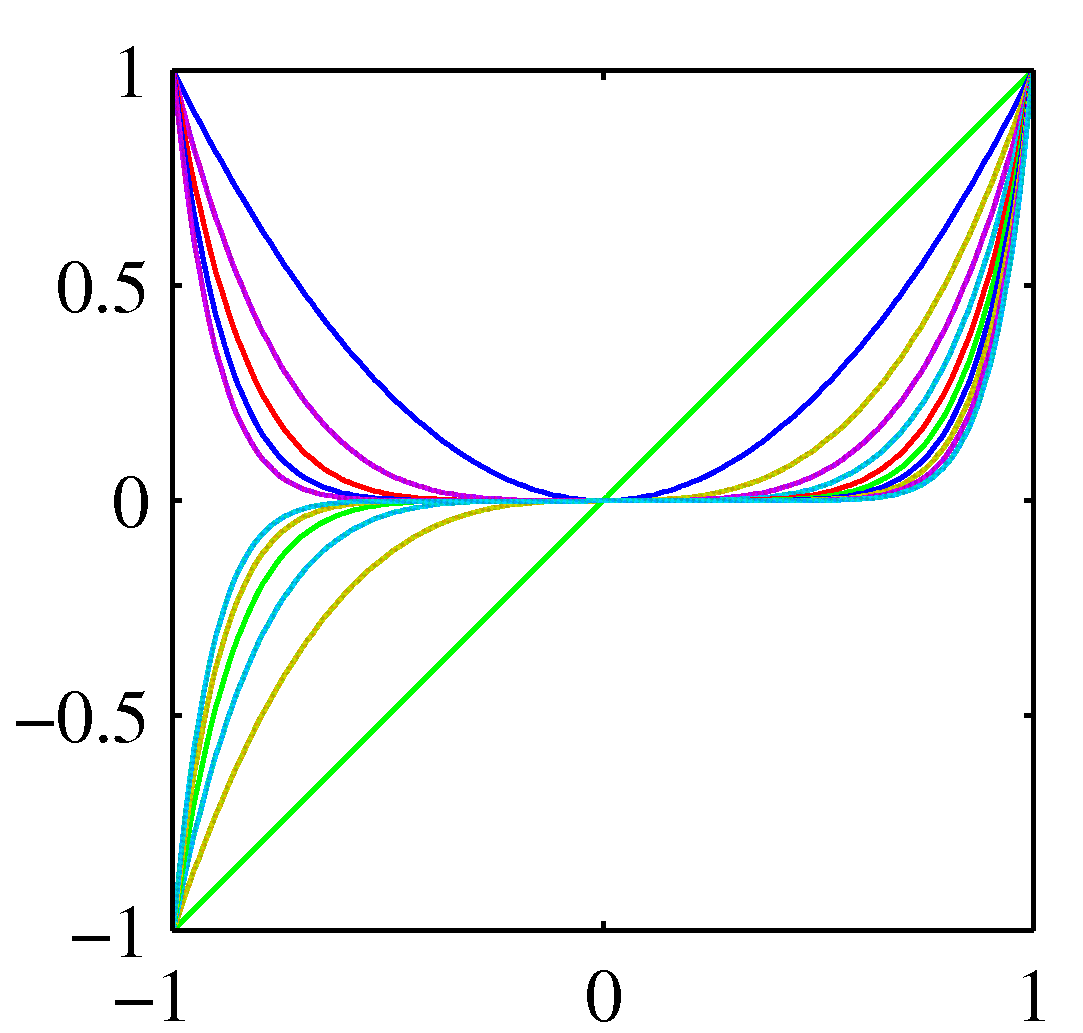
**Different Types of basis Functions**

**Polynomial Basis function**

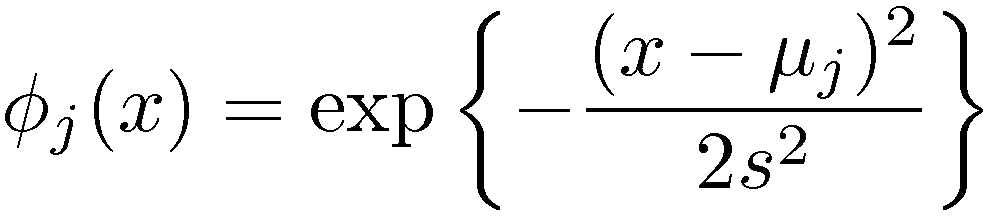


These are global; a small change in x affects all basis functions.

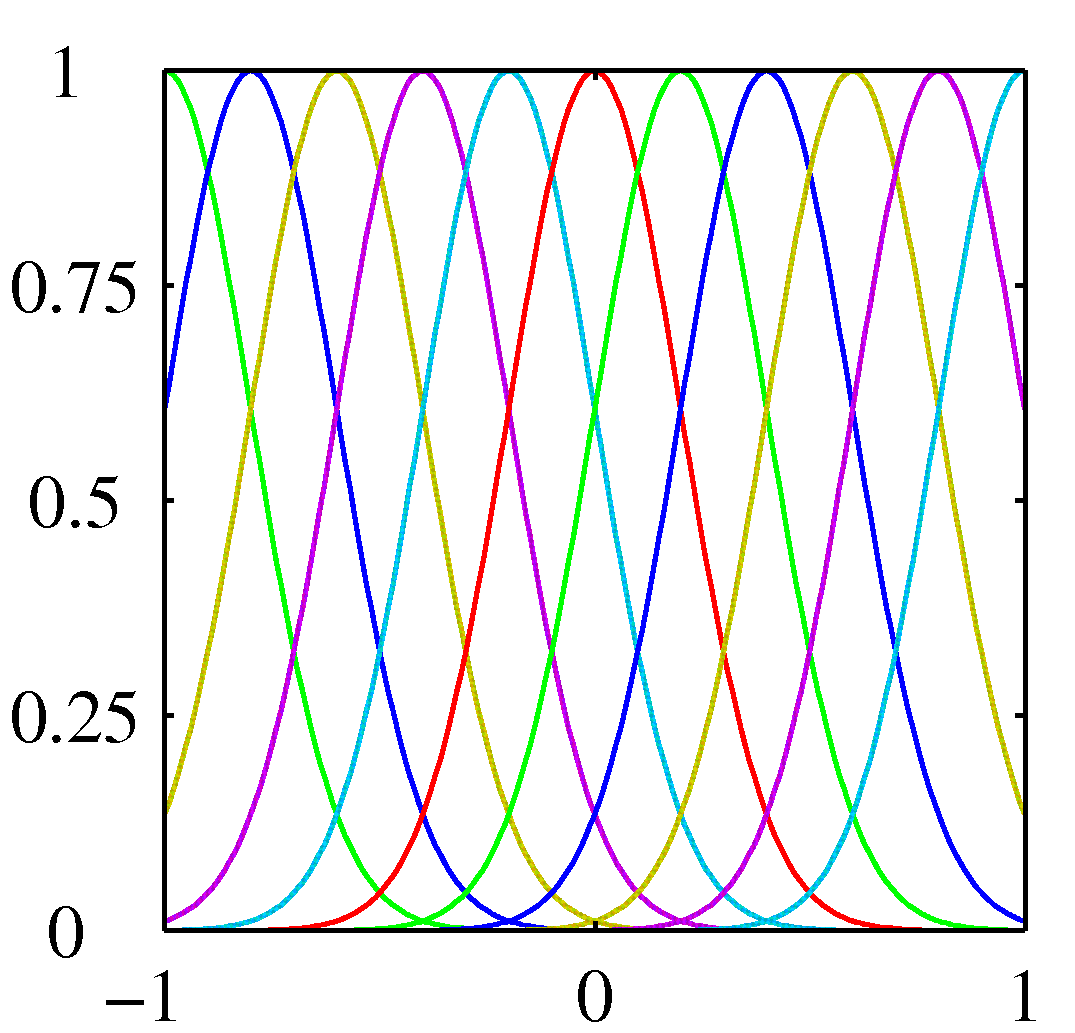




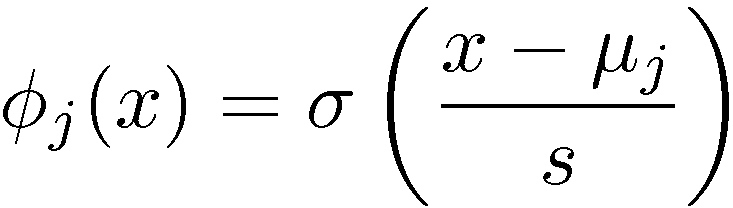
**Gaussian Basis Function:**



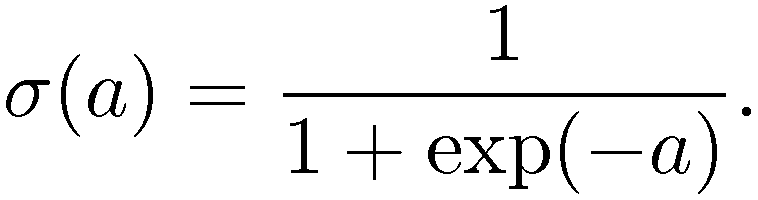
These are local; a small change in x only affect nearby basis functions. And s control location and scale (width).



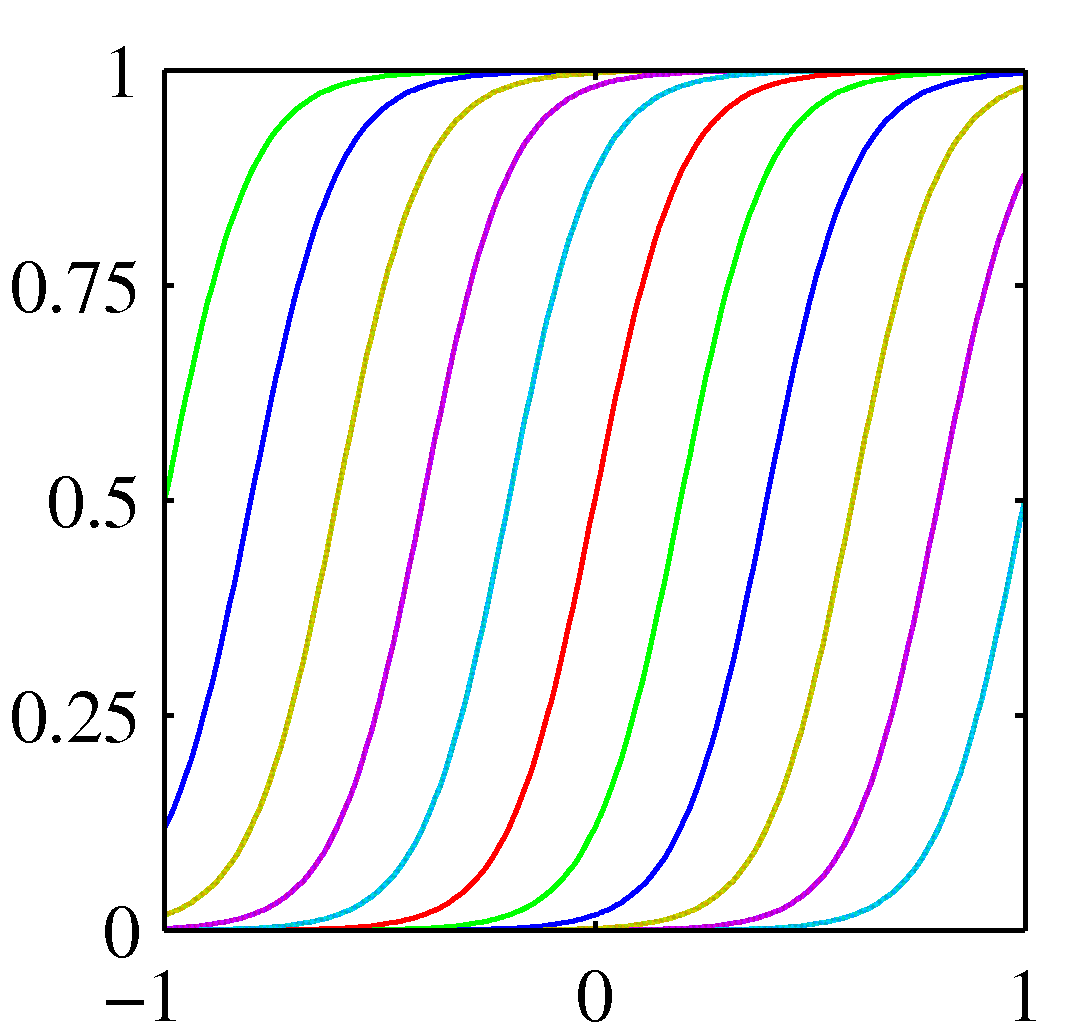
**Sigmoidal basis functions:**



where



Also these are local; a small change in x only affect nearby basis functions. ¹j and s control location and scale (slope).



**Linear regression using different basis functions:**

We want to model data (x1,t1), . . . ,(xN ,tN ), where xi is a vector of D inputs

(predictors) for case i, and ti is the target (response) variable for case i, which is

real-valued.

We are trying to predict t from x, for some future test case, but we are not trying

to model the distribution of x.

Suppose also that we don’t expect the best predictor for t to be a linear function

of x, so ordinary linear regression on the original variables won’t work well.

We need to allow for a non-linear function of x, but we don’t have any theory

that says what form this function should take. What to do?

**An Example Problem**

Use 100 points generated with x uniform from (0, 1) and y set by the formula:

y = sin(x) + noise

**Linear Basis Function Models**

We earlier looked at fitting this data by least-squares linear regression, using not

just x, but also x2, x3, etc., up to (say) x4 as predictors.

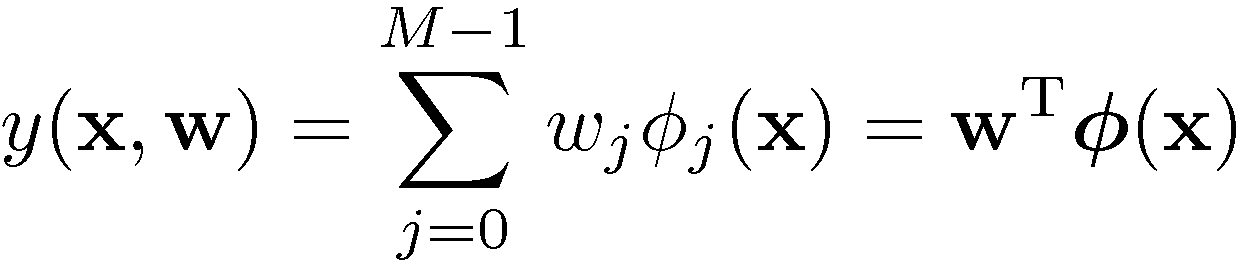
This is an example of a linear basis function model.

In general, we do linear regression of t on φ1(x), φ2(x), . . . , φM−1(x), where the

φj are basis functions, that we have selected to allow for a non-linear function of x.

This gives the following model:

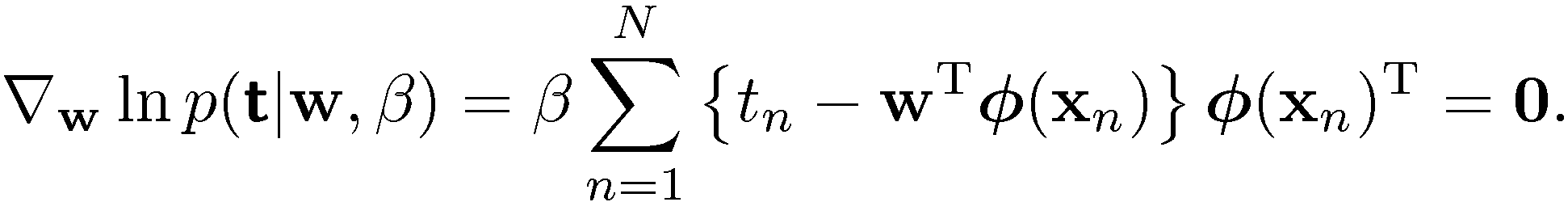
In vector form it can be written as:

 Basically, it is phi(X)\*w.

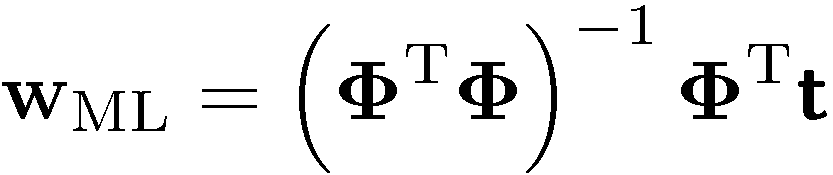
where w is the vector of all M regression coefficients and φ(x) is the vector of all basis function values at input x, including φ0(x) = 1 for the intercept.

**Least Squares Estimation**

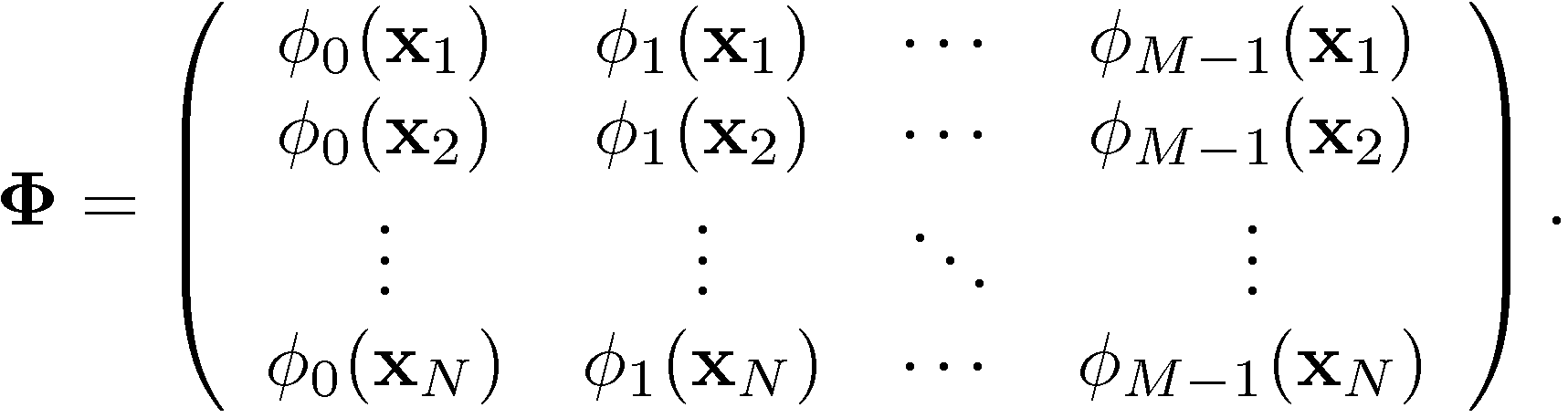
To minimize squared error, its gradient is set to zero and then the equation is solved for w.



Solving it for W, we get:



where phi is:



**Exercise**

1. Compare different basis function’s (polynomial, Gaussian, sigmoidal) by varying their parameters and plotting them.
2. Using different basis functions fit the function y=sin(x)+noise and analyse it’s error (sum of squared error) by varying parameters of the basis function.
3. Try out linear regression problems having multiple input variables xi (i=1,2,…,n). We will do it in the next lab.