

STATISTICAL PATTERN RECOGNITION

LINEAR REGRESSION

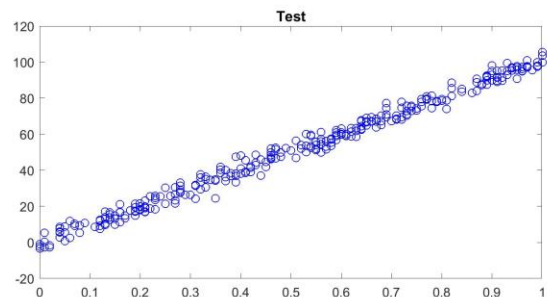
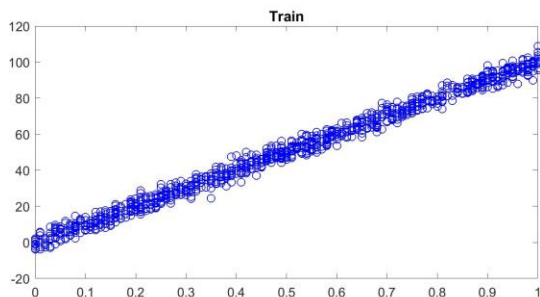
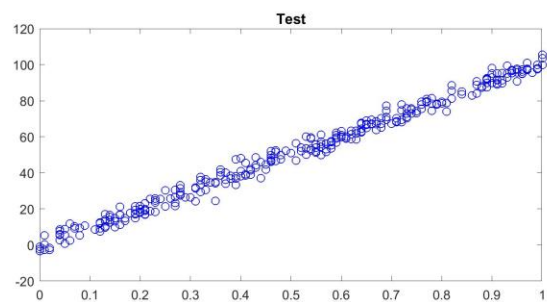
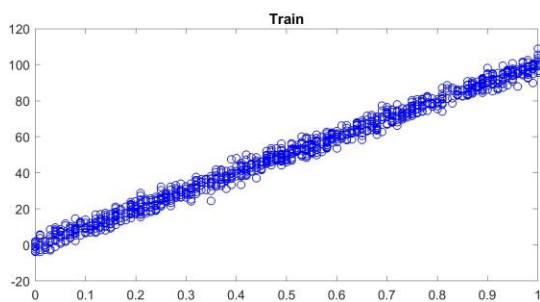
: Preprocessing

نرمالیزه کردن دیتا به بازه ی 0 و 1

$$\mathbf{X}_{\text{norm}} = \frac{(\mathbf{X} - \min(\mathbf{X}))}{(\max(\mathbf{X}) - \min(\mathbf{X}))}$$

```
x1=train(:,1); y1=train(:,2);  
x2=test(:,1); y2=test(:,2);
```

```
x1=(x1-min(x1))./(max(x1)-min(x1));  
x2=(x2-min(x2))./(max(x2)-min(x2));
```



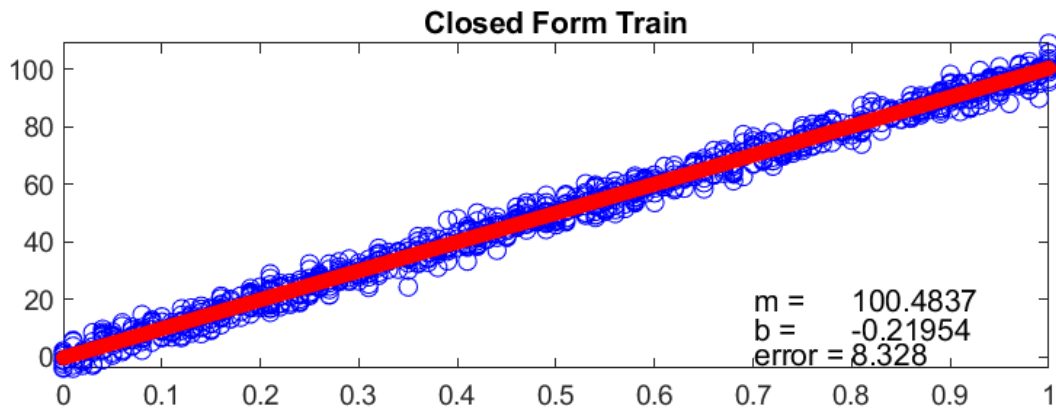
: Closed Form

پیدا کردن نتایج بر اساس فرمول Least Squared Method

$$\theta = (X^T X)^{-1} X^T Y$$

```
x1=[ones(1000,1) x1];
t=(inv(x1'*x1))*x1'*y1;
t0=t(1); t1=t(2);
Y1=t0 + t1.*x1(:,2);
error1=sum((y1-Y1).^2)/1000;
```

%Predicting

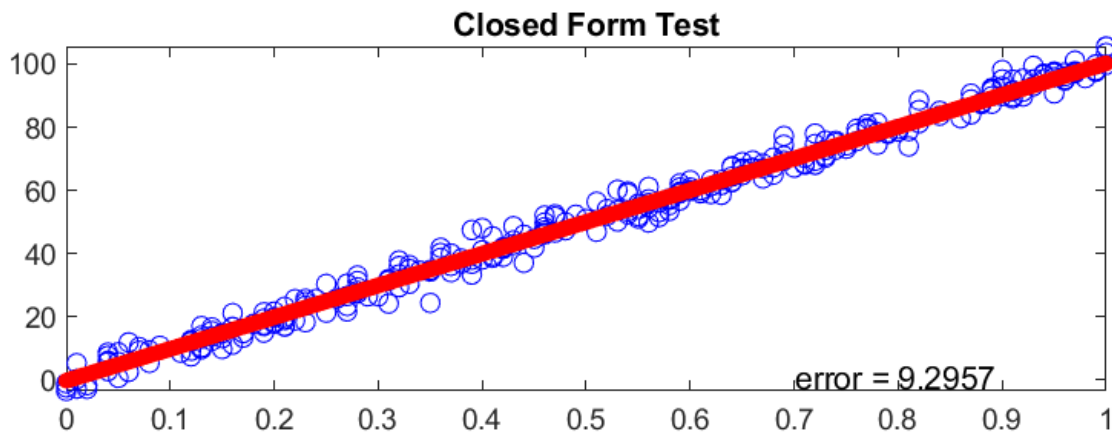


پیش‌بینی با داده‌های تست و محاسبه خطا

$$\hat{Y} = h_{\theta}(X)$$

$$E = \frac{1}{2} \sum (Y - \hat{Y})^2$$

```
Y2=t0 + t1.*x2;
error2=sum((y2-Y2).^2)/300;
```



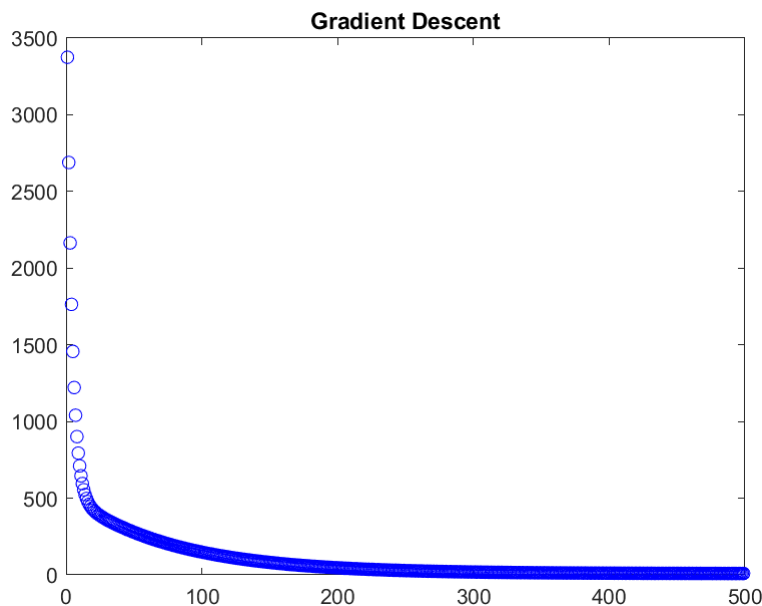
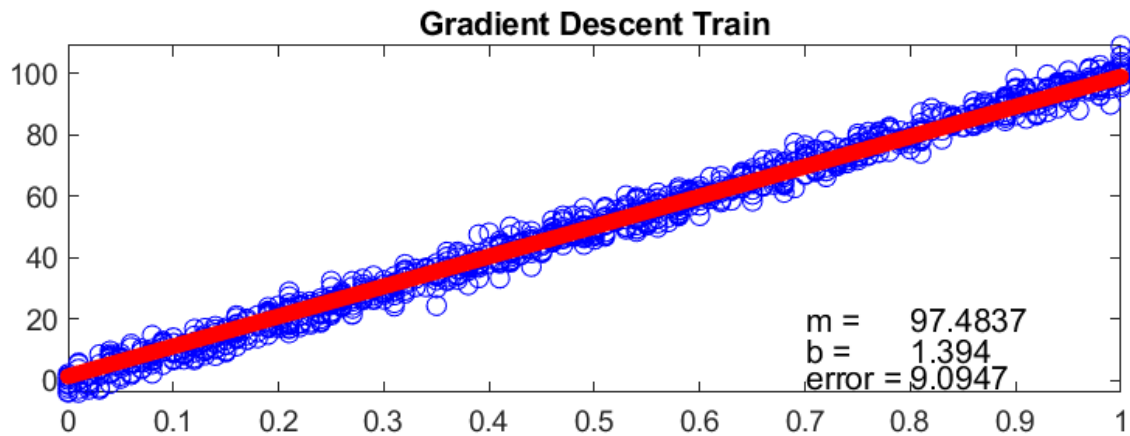
: Batch Gradient Descent

مینیم کردن خطا با مقدار 0 و آپدیت کردن تتاها (بایاس و شیب)

$$\theta_i = \theta_i - \alpha \sum_{j=1}^m (h_{\theta}(\mathbf{x}^{(j)}) - y^{(j)}) x_i^{(j)}$$

```
while error >= tol
    y_predict = X_train*theta';
    error = sum((y_predict-y_train).^2)/1000;
    MSE(ii) = error;

    gradients = (X_train')*(X_train*theta'-y_train);
    theta = theta-(alpha*gradients).';
    b = theta(1);
    m = theta(2);
    ii=ii+1;
end
```



پیشبینی با داده های تست و محاسبه خطا

$$\hat{Y} = \theta_0 + \theta_1 X = b + m X$$

$$E = \frac{1}{2} \sum (Y - \hat{Y})^2$$

```
Y2=m*x2+b;  
error2=sum((y2-Y2).^2)/300;
```

