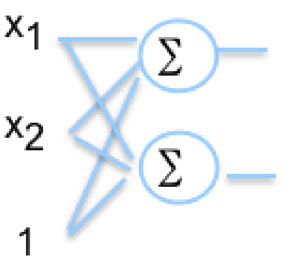
1. Refer to the following code

```
function a=gradient_descent(x,y)
max_loop=2000;
[N,d]=size(x);
X=[x ones(N,1)];
a=rand(1,d+1); hc=0; c=0.1;e=y-X*a';
E=100; loop=1;
while ~hc
   G=mean(X.*(e*ones(1,d+1)));
   a_new=a-c*G;
   y_hat=X*a_new';
   e_new=y_hat-y;
   E_new=mean(e_new.^2);
   if mod(loop,100)==0
   fprintf('loop %d  mse %f\n',loop,E_new);
   end
   if E_new < E & loop < max_loop
      a=a_new;e=e_new;
      E=E_new;
   else
      hc=1;
   end
   loop=loop+1;
end
```

   A. Figure out the architecture of a linear relation.
   B. Write codes for error calculation.
   C. Write codes for gradient calculation.
   D. Write codes for y_hat calculation.

2. Consider the following linear relations




   A. Let $a_1$ and $a_2$ be columns of matrix a. Let ($x_i$   $b_{ji}$ ) be paired data, where i runs from 1 to n.

$$E(\mathbf{a}) = \frac{1}{2n}\sum_{i=1}^{n}\sum_{j=1}^{2}(\mathbf{x}_i^T \mathbf{a}_j - b_{ji})^2$$

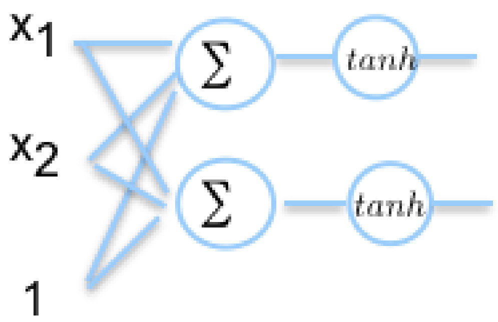
Derive $\frac{d}{d\mathbf{a}}E(\mathbf{a})$

B. Write codes for error calculation.

C. Write codes for gradient calculation.

D. Write codes for y_hat calculation.

3. Consider the following nonlinear relations




A. Let

$$E(a) = \frac{1}{2n}\sum_{i=1}^{n}\sum_{j=1}^{2}\left(\tanh\left(x_i^T a_j\right) - b_{ji}\right)^2$$

Derive $\frac{d}{d\mathbf{a}}E(\mathbf{a})$

B. Write codes for error calculation

C. Write codes for gradient calculation.

D. Write codes for y_hat calculation.

4. Write function gradient_descent to estimate the matrix for linear transformation

```
x=rand(400,2);
z(:,1) = 2*x(:,1)+x(:,2)-1;
z(:,2)=x(:,1)-x(:,2)+1;
a=gradient_descent(x,z)
```

Checked by____________________time__________________

5. Write function gradient_descent to estimate the matrix for nonlinear transformation

```
x=rand(400,2);
z(:,1) = tanh(2*x(:,1)+x(:,2)-1);
z(:,2) = tanh(x(:,1)-x(:,2)+1);
a=gradient_descent(x,z)
```

Checked by______________________time____________________