TP5 : Descente de gradient

Descente de gradient

1. Calculez l'expression analytique de la fonction et sa dérivée.

2. Implémentez l'algorithme DG sous Python pour la fonction E(x).

3. Pour comprendre ce qui fait effectivement la DG, testez l'algorithme implémenté en utilisant des exemples d'exécutions avec des valeurs initiales de x0 et η suivantes :

4. Affichez le minimum trouvé, ainsi que E(xmin) et le nombre d'itérations. Que constatez-vous?

5. Visualisez l'évolution des minimums de la fonction E(x) trouvés au cours des itérations.

6. Testez votre algorithme avec d'autres valeurs de eps et nombre max.

```
00247305519, 4.10499078940642, 4.48300247305519, 4.10499078940642, 4.48300247305519, 4.10499078940642, 4.483
00247305519, 4.10499078940642, 4.48300247305519, 4.10499078940642, 4.48300247305519, 4.10499078940642, 4.483
00247305519]              |
| E(xmin)              |  x**4 - 11.0*x**3 + 41.0*x**2 - 61.0*x + 30.0                |
| nomber des iteration  |  1001                                          |
-----------------------------------------------------------------
|       Pour x0 =  4  et η =  0.17      :                        |
-----------------------------------------------------------------
| xmin                 |  [4, 4.85000000000000, 1.99502000000001, 1.48336482859609, 1.29985340908275, 1.535
20416176282, 1.26598771774773, 1.59966520234823, 1.24239070441334, 1.64867602179201, 1.23755864714305, 1.659
13290291893, 1.23794109228987, 1.65830002269115, 1.23789273162030, 1.65840529160171, 1.23789867346069, 1.658
39235694491, 1.23789794071211, 1.65839395203680, 1.23789803103382, 1.65839375541879, 1.23789801989976, 1.658
39377965613, 1.23789802127225, 1.65839377666839, 1.23789802110307, 1.65839377703668, 1.23789802112392, 1.658
39377699128, 1.23789802112135, 1.65839377699688, 1.23789802112168, 1.65839377699618, 1.23789802112163, 1.658
39377699627, 1.23789802112164, 1.65839377699626, 1.23789802112164, 1.65839377699626, 1.23789802112164, 1.658
39377699626, 1.23789802112164, 1.65839377699626, 1.23789802112164, 1.65839377699626, 1.23789802112164, 1.658
39377699626, 1.23789802112164, 1.65839377699626, 1.23789802112164, 1.65839377699626, 1.23789802112164, 1.658
39377699626]              |
| E(xmin)              |  x**4 - 11.0*x**3 + 41.0*x**2 - 61.0*x + 30.0                |
| nomber des iteration  |  1001                                          |
```

Descente de gradient pour la régression linéaire

1. Calculation des dérivées partielles de la fonction G(a; b) selon a et b

In [11]:
```python
# In our example, G(x,y) = SUM(ax+y - y)
def G(x, y, a, b):
    return np.sum((a*x+b-y)**2)
# define derive of fonction
def derive_G(x, y, a, b):
    return {'dx': 2*(a*x**2+b*x-y*x), 'dy': 2*(a*x+b-y)}

def y(x, a, b):
    return (a*x+b)
```

2. Implémentez l'algorithme DGR

In [12]:
```python
#Gradient Descent implied for a regression linear case with data
def gradient_reg(max_iter, step, x, y, df,eps = 0.01):
    #Const
    nb_sample_data = x.shape[0]
    rg_sample = range(1, nb_sample_data)

    #initialisation
    a = 5 #np.random.random(x.shape[1])
    b = 5 #np.random.random(x.shape[1])
    E = sum([a*x[j]+b-y[j] for j in rg_sample])

#     print(E)
    for i in range(max_iter):
        #Compute gradient
        grad_a = sum([2*(a*x[j]**2+b*x[j]-y[j]*x[j]) for j in rg_sample])
        grad_b = sum([2*(a*x[j]+b-y[j]) for j in rg_sample])

        #Descente de gradient
        a = a - step * grad_a
        b = b - step * grad_b


        #Compute error
        e = sum([a*x[j]+b-y[j] for j in rg_sample])

        if abs(E-e) <= eps:
            return ({'a': a, 'b': b}, i)

        E = e

    return ({'a': a, 'b': b}, max_iter)
```

3. test d'algorithme implémenté

In [13]:
```python
from sklearn import datasets
x_reg, y_reg = datasets.make_regression(n_samples=100, n_features=1, noise=7)
all_final_reg = []
print(x_reg.reshape(100).shape)
print(y_reg.shape)
```

```
(100,)
(100,)
```

In [14]:
```python
def show_res_reg(vect,derive_G):
    print("**************************************************************")
    res,nb_ite = gradient_reg(vect[0],vect[1],vect[2],vect[3],derive_G)
    print("for initial max iteration = ",vect[0]," , step = ",vect[1])
    print(' ==> final_a = ',res['a'],' final_b = ',res['b'],'nb_iteration = ',nb_ite)
    pry = [y(x_d[0],res['a'],res['b']) for x_d in x_reg]
#     print("final result =  ",G(x_reg,y_reg,res['a'],res['b']))
    print('E(', x_reg[0][0], ') = ', pry[0])
    all_final_reg.append(pry)
    print("**************************************************************")
```

4. Affichage des resultas trouvés

In [15]:
```python
all_final_reg = []
show_res_reg([100,0.001,x_reg,y_reg],derive_G)
show_res_reg([500,0.001,x_reg,y_reg],derive_G)
show_res_reg([1000,0.001,x_reg,y_reg],derive_G)
show_res_reg([100,0.01,x_reg,y_reg],derive_G)
show_res_reg([100,1,x_reg,y_reg],derive_G)
```

```
**************************************************************
for initial max iteration =  100  , step =  0.001
 ==> final_a =  [93.66725942]  final_b =  [0.01392354] nb_iteration =  49
E( -0.7493494090740423 ) =  [-70.17558196]
**************************************************************
**************************************************************
for initial max iteration =  500  , step =  0.001
 ==> final_a =  [93.66725942]  final_b =  [0.01392354] nb_iteration =  49
E( -0.7493494090740423 ) =  [-70.17558196]
**************************************************************
**************************************************************
for initial max iteration =  1000  , step =  0.001
 ==> final_a =  [93.66725942]  final_b =  [0.01392354] nb_iteration =  49
E( -0.7493494090740423 ) =  [-70.17558196]
**************************************************************
**************************************************************
for initial max iteration =  100  , step =  0.01
 ==> final_a =  [-2.57176602e+08]  final_b =  [-76344971.72586623] nb_iteration =  100
E( -0.7493494090740423 ) =  [1.16370163e+08]
**************************************************************
**************************************************************
for initial max iteration =  100  , step =  1
 ==> final_a =  [-1.51482288e+235]  final_b =  [-4.49648641e+234] nb_iteration =  100
E( -0.7493494090740423 ) =  [6.85482986e+234]
**************************************************************
```

5. utilisation des stats.linregress de scipy

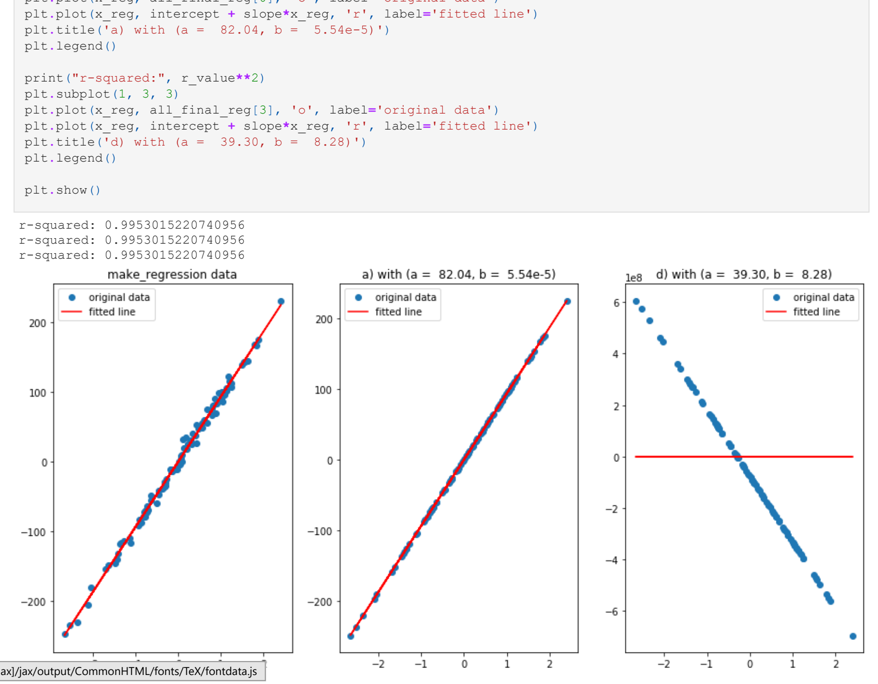
In [16]:
```python
from scipy import stats
slope, intercept, r_value, p_value, std_err = stats.linregress(x_reg.reshape(100), y_reg)
np.array(all_final_reg[0]).shape
```

Out[16]:
```
(100, 1)
```

6. Visualisation

In [17]:
```python
print("r-squared:", r_value**2)

plt.rcParams["figure.figsize"] = (15,7)
plt.subplot(1, 3, 1)
plt.plot(x_reg, y_reg, 'o', label='original data')
plt.plot(x_reg, intercept + slope*x_reg, 'r', label='fitted line')
plt.title('make_regression data')
plt.legend()


print("r-squared:", r_value**2)
plt.subplot(1, 3, 2)
plt.plot(x_reg, all_final_reg[0], 'o', label='original data')
plt.plot(x_reg, intercept + slope*x_reg, 'r', label='fitted line')
plt.title('a) with (a =  82.04, b =  5.54e-5)')
plt.legend()

print("r-squared:", r_value**2)
plt.subplot(1, 3, 3)
plt.plot(x_reg, all_final_reg[3], 'o', label='original data')
plt.plot(x_reg, intercept + slope*x_reg, 'r', label='fitted line')
plt.title('d) with (a =  39.30, b =  8.28)')
plt.legend()

plt.show()
```

```
r-squared: 0.9953015220740956
r-squared: 0.9953015220740956
r-squared: 0.9953015220740956
```



Links

- E-mail : zakaria.abbou199434@gmail.com
- GitHub : github.com/ZakariaAABBOU
- Linkedin : linkedin.com/in/zakaria-aabbou/