Task 1:
Для каждой из следующих функций $f(x)$ выписать первый дифференциал. Для скалярных функций векторного аргумента найти вектор-градиент $\nabla f(x)$. Для скалярных функций матричного аргумента найти матрицу-градиент $\nabla f(X)$.

2. $f(x)=\exp(A\exp(B\exp(Cx)))$, где $\exp(y)=[\exp(y1),\dots,\exp(yn)]T$, $x\in\mathbb{R}n$
8. $f(X)=\det(XTAX)$, $X\in\mathbb{R}n\times n$

Solution:
$f(x)=\exp(A\exp(B\exp(Cx)))$, где $\exp(y)=[\exp(y1),\dots,\exp(yn)]T$, $x\in\mathbb{R}n$
$df(x;dx)=d(\exp(A\exp(B\exp(Cx))))=\{d(\exp(x))=dx\exp(x)\}=\exp(A\exp(B\exp(Cx)))$ $d(A\exp(B\exp(Cx)))=\exp(A\exp(B\exp(Cx)))$ $A\exp(B\exp(Cx))$ $d(B\exp(Cx))=\exp(A\exp(B\exp(Cx)))$ $A\exp(B\exp(Cx))$ $B\exp(Cx)$ $d(Cx)$ $=ABC\exp(A\exp(B\exp(Cx)))$ $\exp(B\exp(Cx))$ $\exp(Cx)$ dx .

$\nabla f(x)=(ABC\exp(A\exp(B\exp(Cx)))$ $\exp(B\exp(Cx))$ $\exp(Cx))T$.

8. $f(X)=\det(XTAX)$, $X\in\mathbb{R}n\times n$
First Derivative:
$d(det(X))=det(X)\langle X-T,dX\rangle$ – из таблицы стандартных производных
$d(\det(XTAX))=\det(XTAX)\langle(XTAX)-T,d(XTAX)\rangle=$ $\det(XTAX)\langle X-1A-TX-T,d(XT)AX+$ $XTAdX\rangle$
Чтобы найти градиент, приведем $df(X)$ к канонической форме $df(X)=$ $\langle\nabla f(X),dX\rangle$ $df(X)=$ $\det(XTAX)(\langle X-1A-TX-T,d(XT)AX\rangle+$ $\langle X-1A-TX-T,XTAdX\rangle)=$
$\det(XTAX)(\langle X-1A-TX-TXTAT,d(XT)\rangle+$ $\langle ATXX-1A-TX-T,dX\rangle)=$ $\det(XTAX)(\langle X-1,d(XT)\rangle+$ $\langle X-T,dX\rangle)=$ $\det(XTAX)(\langle X-T,dX\rangle+$ $\langle X-T,dX\rangle)$ $=$ $\det(XTAX)\langle X-T+$ $X-T,dX\rangle=$ $\langle\det(XTAX)(2X-T),dX\rangle$

$\nabla f(X)=$ $\det(XTAX)(2X-T)$

Task 4:
1. $\max x\in\mathbb{R}+n,1Tx=1aTx$, где $a\in\mathbb{R}+n$
$\max aTx$ $\leq$ $\max(aT)\max(x)$ $\max aTx$ $=$ $\max(aT)$

2. $\min x\in\mathbb{R}n\mathrm{vec}(xaT)T\mathrm{vec}(xbT)-cTx$, где $a,b\in\mathbb{R}+n$ и $\mathrm{vec}(\cdot)$ обозначает векторизацию.
$\mathrm{vec}(xaT)T\mathrm{vec}(xbT)-cTx=tr((xaT)T(xbT))-cTx=tr(axTxbT)-cTx$ $d(tr(axTxbT)-cTx)=$ $d(tr(axTxbT))-d(cTx)=$ $tr(d(axTxbT))-cTdx=$ $tr(d(axT)xbT+$ $axTd(xbT))$ $-cTdx=$ $tr(d(axT)xbT)+tr($ $axTd(xbT))-cTdx=tr(bxTdxaT)+tr($ $axTdxbT)-cTdx=$ $tr(aTbxTdx)+tr($ $bTaxTdx)-cTdx$ $\nabla f(x)=$ $aTbxT+bTaxT-cT$
$\nabla f(x)=0$ $aTbxT+bTaxT-cT=0$ $(aTb+bTa)xT=$ cT $xT=$ $(aTb+bTa)-1cT$ $x=$ $(aTb+bTa)-Tc$
$\min x\in\mathbb{R}n\mathrm{vec}(xaT)T\mathrm{vec}(xbT)-cTx=$ $tr(a(aTb+bTa)-1cT((aTb+bTa)-Tc)bT)-cT((aTb+bTa)-Tc)$

3. $\min X\in\mathbb{R}n\times n\mathrm{vec}(X)T\mathrm{vec}(AXB)+\mathrm{tr}(ATX)$, где $A,B\succ0$ (имеют положительные собственные числа).

vec(X)Tvec(AXB)+tr(ATX)=tr((X)T(AXB))+tr(ATX)=tr($BXTAX$)+tr(ATX)
d(tr($BXTAX$)+tr(ATX))=d(tr($BXTAX$))+d(tr(ATX))=tr(d($BXTAX$))+tr(d(ATX))=tr(Bd(XT)AX+$BXTAd$(X))+tr(ATd(X))=tr($XTATd$(X)BT)+tr($BXTAd$(X))+tr(ATd(X))=
tr($BTXTATd$(X))+tr($BXTAd$(X))+tr(ATd(X))=tr(($BTXTAT$+$BXTA$+AT)d(X)) $\nabla f(X)=AXB+ATXTBT+A$
$\nabla f(X)=0$ $AXB+ATXTBT+A=0$
т.к A,B>0 , то решений нет

Task4:

1. convolve(x) :

```
def f_c(x):
    ret = np.zeros((dim,))
    ret[:-1] += x[1:]; ret += -2*x; ret[1:] += x[:-1]
return ret
```

$$f\_C\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}\right) = \begin{bmatrix} -2x_1 + x_2 \\ x_3 - 2x_2 + x_1 \\ x_4 - 2x_3 + x_2 \\ -2x_4 + x_{23} \end{bmatrix}$$

2. Jacobian for f_c(x)

$$J(x) = \begin{bmatrix} -2 & 1 & 0 & 0 \\ 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 1 & -2 \end{bmatrix}$$

3. VJP

$$VJP(v;x) = v^T J(x) = \begin{bmatrix} v_1 & v_2 & v_3 & v_4 \end{bmatrix} \begin{bmatrix} -2 & 1 & 0 & 0 \\ 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 1 & -2 \end{bmatrix} =$$

$$= \begin{bmatrix} -2v_1 + v_2 \\ v_1 - 2v_2 + v_3 \\ v_2 - 2v_3 + v_4 \\ v_3 - 2v_4 \end{bmatrix}^T = \begin{bmatrix} -2v_1 + v_2 \\ v_3 - 2v_2 + v_1 \\ v_4 - 2v_3 + v_3 \\ -2v_4 + v_3 \end{bmatrix}^T$$

= VJP (v;x) = (f_c(v))^T = (convolve(v))^T

gradient of initial function

$\mathrm{grad}f(x) = J_2\,(\mathrm{convolve}(f_1(x)))J_c\,(f1(x))J_1(x) = VJPc\,(BJP_2\,(1.0;\mathrm{convolve}(f_1(x)));f_1(x))J_1(x) = VJP_c(\mathrm{grad}f_2(\mathrm{convolve}(f_1(x)))\text{^}T\,;\,f1(x))J_1(x)$