Hints for Exercise set 11

Problem 1: Use the representation of $\mathcal{X}(i_1,\dots,i_d) = \square \cdot \square \cdot \square \cdots \square \cdot \square \quad (*)$

How can you write the sum of two expressions like $(*)$ in the same form as $(*)$ (maybe with matrices/vectors of double size)?

Problem 2: Recall that you can write $X^{(\mu)} = U_\mu^{(2)} \cdot \left(X_{<\mu}^{<\mu-1>} \otimes X_{>\mu}^{<1>} \right)$ for lexicographical multiindices.

what is the size (therefore an upper bound on the rank) of this?

Problem 4: For the first part, write

$$\operatorname{vec}(\mathcal{X}) = (I \otimes \dots \otimes I \otimes U_1^L)(\dots)(I \otimes U_{d-1}^L) \cdot U_d^L$$

and use that $U_1^L, \dots, U_{d-1}^L$ have orthonormal columns.

Problem 5: Rewrite the average of the entries of $\mathcal{X}$ as the inner product between two suitable tensors? Then you can use Problem 5 from last week!

Solutions to Exercise Set 11

Problem 1

Let $\mathcal{X}(i_1,\dots,i_d) = \sum_{k_1,\dots,k_{d-1}} U_1(1,i_1,k_1)U_2(k_1,i_2,k_2)\cdots U_d(k_{d-1},i_d,1)$ ranks $r_1,\dots,r_{d-1}$

and $\mathcal{Y}(i_1,\dots,i_d) = \sum_{h_1,\dots,h_{d-1}} V_1(1,i_1,h_1)V_2(h_1,i_2,h_2)\cdots V_d(h_{d-1},i_d,1)$ ranks $s_1,\dots,s_{d-1}$

$\mathcal{Z} := \mathcal{X} + \mathcal{Y}$

$$\mathcal{Z}(i_1,\dots,i_d) = \begin{bmatrix} U_1(i_1) & V_1(i_1)\end{bmatrix} \cdot \begin{bmatrix} U_2(i_2) & 0 \\ 0 & V_2(i_2)\end{bmatrix}\begin{bmatrix} U_3(i_3) & 0 \\ 0 & V_3(i_3)\end{bmatrix}\cdots\begin{bmatrix} U_d(i_d) \\ V_d(i_d)\end{bmatrix}$$

More formally, $\mathcal{Z}(i_1,\dots,i_d) = \sum_{\alpha_1=1}^{r_1+s_1}\cdots\sum_{\alpha_{d-1}=1}^{r_{d-1}+s_{d-1}} W_1(1,i_1,\alpha_1)W_2(\alpha_1,i_2,\alpha_2)\cdots W_d(\alpha_{d-1},i_d,1)$

where $W_\ell(\alpha_{\ell-1},i_\ell,\alpha_\ell) = \begin{cases} U_\ell(\alpha_{\ell-1},i_\ell,\alpha_\ell) & \text{if } \alpha_{\ell-1}\le r_{\ell-1} \text{ and } \alpha_\ell \le r_\ell \\ V_\ell(\alpha_{\ell-1}-r_{\ell-1},i_\ell,\alpha_\ell - r_\ell) & \text{if } \alpha_{\ell-1} > r_{\ell-1} \text{ and } \alpha_\ell > r_\ell \\ 0 & \text{otherwise.}\end{cases}$

Problem 2

Let $\mathcal{X}(i_1,\dots,i_d) = \sum_{\alpha_1=1}^{r_1}\cdots\sum_{\alpha_{d-1}=1}^{r_{d-1}} U_1(1,i_1,\alpha_1)U_2(\alpha_1,i_2,\alpha_2)\cdots U_d(\alpha_{d-1},i_d,1)$

be a d-dimensional tensor in TT-decomp. with ranks $(r_1,\dots,r_{d-1})$.

Then we can write

$$X^{(\mu)} = U_\mu^{(2)} \cdot \left(X_{<\mu}^{(\mu-1)} \otimes X_{>\mu}^{(1)}\right)$$ for lexicographical multi-indices

this has size $n_\mu \times (r_{\mu-1}\cdot r_\mu)$

$\Rightarrow \operatorname{rank}(X^{(\mu)}) \le r_{\mu-1}\cdot r_\mu$ for $\mu = 1,\dots,d-1$
(with the convention that $\mu_0 = 1$).

Problem 4

a) If the TT decomposition is left-orthogonal, we have that

$\|\mathcal{X}\| = \|\operatorname{vec}(\mathcal{X})\|_2 = \|(I\otimes\cdots\otimes I\otimes U_1^L)\cdot\ldots\cdot(I\otimes U_{d-1}^L)\cdot U_d^L\|_2$ ⊛

By left-orthogonality, for $\mu = 1,\dots,d-1$ we have $(U_\mu^L)^T\cdot U_\mu^L = I$, so the matrices $I\otimes\cdots\otimes I\otimes U_\mu^L$ have orthonormal columns $\mu = 1,\dots,d-1$. Therefore, ⊛ $= \|U_d^L\|_2 =$
$= \|\operatorname{vec}(U_d^L)\|_2 = \|U_d\|_2$.

By left-orthogonality, $X_{\le\mu}^T \cdot X_{\le\mu} = I$ for $\mu = 1, \dots, d-1$.

As $X^{<\mu>} = \underbrace{X_{\le\mu}}_{\text{orthogonal}} \cdot X_{\ge\mu+1}$, the sing. values of $X^{<d-1>}$ are the same as $X_{\ge d} = U_d$.

b) Analogous to the previous point, using $\|\mathcal{X}\| = \| U_1^R (U_2^R \otimes I) \cdots (U_d^R \otimes I \otimes \cdots \otimes I) \|_2$ and the fact that $U_\mu^R \cdot (U_\mu^R)^T = I$ for $\mu = 2, \dots, d$.

Problem 5

If $\mathcal{X}$ has ranks $(R, \dots, R)$ the complexity of computing the average is $O(R^2(n_1 + \dots + n_d))$

Indeed, we can write $\bar{\mathcal{X}} = \frac{1}{n_1 n_2 \cdots n_d} \langle \mathcal{X}, \mathcal{E} \rangle$ where $\mathcal{E} \in \mathbb{R}^{n_1 \times \dots \times n_d}$ has all entries equal to 1; this $\mathcal{E}$ can be written in TT-format with ranks $(1, \dots, 1)$ (with $U_\mu \in \mathbb{R}^{1 \times n_\mu \times 1}$ vector of all ones) so by Problem 5 in Exercise set 10 we have that the complexity of the inner product is $O(R^2(n_1 + \dots + n_d))$.

```
% Problem 3
function T = tt_tensor(U)
    % creating a full tensor out of TT cores U
    d = length(U);
    n = zeros(d,1);
    r = zeros(d+1,1);
    r(end) = 1;
    for mu = 1:d
        n(mu) = size(U{mu}, 2);
        r(mu) = size(U{mu}, 1);
    end
    T = U{1};
    for mu = 2:d
        T = tenmat(T, ndims(T))' * tenmat(U{mu}, 1);
        T = reshape(T, [n(1:mu)', r(mu+1)]);
    end
end

function Ucore = tt_svd(X, rnk, tol)
    d = ndims(X); %order of tensor
    n = size(X);
    r = zeros(d+1, 1);
    r(1) = 1;
    r(end) = 1;
    if tol == 0
        r(2:end-1) = rnk;
    else
        tol = tol / sqrt(d-1);
    end

    Ucore = cell(d);
    for mu = 1:d-1
        X2 = reshape(X, r(mu) * n(mu), prod(n(mu+1:end)));
        [U, S, V] = svd(X2);
        if tol == 0
            U = U(:, 1:r(mu+1));
            S = S(1:r(mu+1), 1:r(mu+1));
            V=V(:,1:r(mu+1));
        else
            K = find(diag(S)>tol);
            U = U(:,K);
            S = S(K, K);
            V = V(:, K);
            r(mu + 1) = size(U, 2);
        end
        Ucore{mu} = reshape(U, [r(mu), n(mu), r(mu+1)]);
        X = reshape(S*V', [r(mu+1), n(mu+1:end)]);
    end
    Ucore{d}=X;
end

% Problem 4
```

```
X = rand(6, 5, 4, 3);
d = ndims(X);
U = tt_svd(X, [], 1e-8);

% left orthogonalization
for mu = 1:d-1
    UL = tenmat(U{mu}, 3)';
    [Q, R] = qr(UL, 0);
    U{mu} = reshape(Q, size(U{mu}));
    U{mu+1} = ttm(U{mu+1}, R, 1);
end
T = tt_tensor(U);
norm(X(:) - T(:))

U = tt_svd(X, [], 1e-8);
% right orthogonalization
for mu = d:-1:2
    UR = tenmat(U{mu}, 1);
    [Q, R] = qr(UR', 0);
    U{mu} = reshape(Q', size(U{mu}));
    U{mu-1} = ttm(U{mu-1}, R', 3);
end
T = tt_tensor(U);
norm(X(:) - T(:))

function Y = ttm(X, A, n)
    Y = tenmat(X, n);
    Y = A * Y;
    sz = size(X);
    sz2 = [size(A,1), sz(1:(n-1)), sz((n+1):end)];
    Y = reshape(Y, sz2);
    perm = [2:n, 1, (n+1):length(sz)];
    Y = permute(Y, perm);
end

function Xn = tenmat(X, n)
    N = ndims(X);
    sz = size(X);
    m = setdiff(1:N, n);
    X = permute(X,[n m]);
    Xn = reshape(X, sz(n), prod(sz(m)));
end
```

Published with MATLAB® R2020a