University of Oxford

Nonlinear Analysis and Applications -Problem Sheet Three

Aili Shao

Throughout this sheet, we assume that $p \in (1, \infty)$ and that Ω is a proper, bounded Lipschitz domain in $\mathbb{R}^n$.

Problem 1. *Let $\frac{1}{p} + \frac{1}{p'} = 1$ and assume that $u_j \rightharpoonup u$ in $W^{1,p}(\Omega)$ and $\sigma_j \rightharpoonup \sigma$ in $L^{p'}(\Omega, \mathbb{R}^n)$, where $\operatorname{div}\sigma_j = 0$ in $\mathcal{D}'(\Omega)$. Show that*

$$\nabla u_j \cdot \sigma_j \overset{*}{\rightharpoonup} \nabla u \cdot \sigma \text{ in } C_0(\Omega)^*.$$

Proof. We first consider for $\varphi \in C_c^\infty(\Omega)$,

$$\begin{aligned}
\left|\int_\Omega \nabla u_j \cdot \sigma_j \varphi \,\mathrm{d}x - \int_\Omega \nabla u \cdot \sigma \varphi \,\mathrm{d}x\right| &= \left|\int_\Omega (\nabla u_j - \nabla u) \cdot \sigma_j \varphi \,\mathrm{d}x + \int_\Omega (\sigma_j - \sigma) \nabla u \varphi \,\mathrm{d}x\right| \\
&\leq \left|\int_\Omega (\nabla u_j - \nabla u) \cdot \sigma_j \varphi \,\mathrm{d}x\right| + \left|\int_\Omega (\sigma_j - \sigma) \nabla u \varphi \,\mathrm{d}x\right| \\
&= \left|-\int_\Omega (u_j - u) \operatorname{div}(\sigma_j \varphi) \,\mathrm{d}x\right| + \left|\int_\Omega (\sigma_j - \sigma) \nabla u \varphi \,\mathrm{d}x\right| \\
&= \left|-\int_\Omega (u_j - u)(\sigma_j \cdot \nabla \varphi) \,\mathrm{d}x\right| + \left|\int_\Omega (\sigma_j - \sigma) \nabla u \varphi \,\mathrm{d}x\right|
\end{aligned}$$

where the last line follows from the fact that $\operatorname{div}(\sigma_j \varphi) = \operatorname{div}\sigma_j \cdot \varphi + \sigma_j \cdot \nabla \varphi = \sigma_j \cdot \nabla \varphi$ provided that $\operatorname{div}\sigma_j = 0$ in $\mathcal{D}'(\Omega)$. Since $u \in W^{1,p}(\Omega)$, $\nabla u \in L^p(\Omega)$, and $\varphi \nabla u \in L^p(\Omega)$, the second term converges to 0 since $\sigma_j \rightharpoonup \sigma$ in $L^{p'}$ with $\frac{1}{p} + \frac{1}{p'} = 1$. As for the first term,

$$\begin{aligned}
\left|-\int_\Omega (u_j - u)(\sigma_j \nabla \varphi) \,\mathrm{d}x\right| &\leq \int_\Omega |(u_j - u)\sigma_j \cdot \nabla \varphi| \,\mathrm{d}x \\
&\leq \|u_j - u\|_{L^p(\Omega)} \|\sigma_j\|_{L^{p'}(\Omega)} \|\nabla \varphi\|_{L^\infty(\Omega)} \\
&\to 0 \text{ as } j \to \infty
\end{aligned}$$

since $W^{1,p}(\Omega)$ is compactly embedded into $L^p(\Omega)$ which in turn implies that $u_j \to u$ in $L^p(\Omega)$.
This shows that

$$\int_\Omega \nabla u_j \cdot \sigma_j \varphi \, \mathrm{d}x \to \int_\Omega \nabla u \cdot \sigma \varphi \, \mathrm{d}x$$

as $j \to \infty$ for all $\varphi \in C_c^\infty(\Omega)$. Note that for each $\varphi \in C_0(\Omega)$, we can find a sequence $\varphi_k \in C_c^\infty$ such that $\varphi_k \to \varphi$ in $C_0(\Omega)$. Then we can conclude that

$$\int_\Omega \nabla u_j \cdot \sigma_j \varphi \, \mathrm{d}x \to \int_\Omega \nabla u \cdot \sigma \varphi \, \mathrm{d}x$$

as $j \to \infty$ for all $\varphi \in C_0(\Omega)$. □

Using Piola's identity deduce that if $v_j \rightharpoonup v$ in $W^{1,n}(\Omega, \mathbb{R}^n)$, then

$$\det \nabla v_j \overset{*}{\rightharpoonup} \det \nabla v \text{ in } C_0(\Omega)^*.$$

Proof. First note that by n-homogeneity of det, we have

$$n \det(\nabla v_j) = \nabla v_j : \operatorname{cof}(\nabla v_j).$$

For $\varphi \in C_c^\infty$, we have

$$\begin{aligned}
\int_\Omega (\det(\nabla v_j) - \det(\nabla v))\, \varphi \, \mathrm{d}x =& \frac{1}{n} \int_\Omega (\nabla v_j : \operatorname{cof}(\nabla v_j) - \nabla v : \operatorname{cof}\nabla v)\, \varphi \, \mathrm{d}x \\
=& \frac{1}{n} \int_\Omega \left[\sum_{k,i=1}^n (\nabla v_j)_{ik} (\operatorname{cof}\nabla v_j)_{ik} - \sum_{k,i=1}^n (\nabla v)_{ik} (\operatorname{cof}\nabla v)_{ik} \right] \varphi \, \mathrm{d}x.
\end{aligned}$$

Consider the ith row only, that is

$$\frac{1}{n} \int_\Omega [(\nabla v_j)_i \cdot (\operatorname{cof}\nabla v_j)_i - (\nabla v)_i \cdot (\operatorname{cof}\nabla v)_i]\, \varphi \, \mathrm{d}x.$$

Note that $v_j \rightharpoonup v$ weakly in $W^{1,n}(\Omega, \mathbb{R}^n)$. Then $(v_j)_i \rightharpoonup (v)_i$ in $W^{1,n}(\Omega)$. By Piola's identity, we know that $\operatorname{div}(\operatorname{cof}\nabla v_j) = 0$ in the distributional sense for $v_j \rightharpoonup\in W^{1,n}(\Omega, \mathbb{R}^n) \hookrightarrow W^{1,n-1}_{\mathrm{loc}}(\Omega, \mathbb{R}^n)$. It is sufficient to show that $\sigma_j := (\operatorname{cof}\nabla v_j)_i \rightharpoonup (\operatorname{cof}\nabla v)_i$ in $L^{\frac{n}{n-1}}(\Omega, \mathbb{R}^n)$ for $v_j \in W^{1,n}(\Omega, \mathbb{R}^n)$. We prove it by induction. When $n = 2$, consider $v_j = (v_j^1, v_j^2)^T \in W^{1,}(\Omega, \mathbb{R}^2)$, then $\nabla v_j = \begin{bmatrix} \frac{\partial v_j^1}{\partial x_1} & \frac{\partial v_j^1}{\partial x_2} \\ \frac{\partial v_j^2}{\partial x_1} & \frac{\partial v_j^2}{\partial x_2} \end{bmatrix}$ and $\operatorname{cof}(\nabla v_j) = \begin{bmatrix} \frac{\partial v_j^2}{\partial x_2} & -\frac{\partial v_j^2}{\partial x_1} \\ -\frac{\partial v_j^1}{\partial x_2} & \frac{\partial v_j^1}{\partial x_1} \end{bmatrix}$.
It follows from $(\nabla v_j)_i \rightharpoonup (\nabla v)_i$ in $L^2(\Omega, \mathbb{R}^2)$ that $(\operatorname{cof}(\nabla v_j)_i \rightharpoonup \operatorname{cof}(\nabla v)_i$ in $L^2(\Omega, \mathbb{R}^n)$. The statement is true for $n = 2$. Now assume that the statement is true for $n-1$, that is, if $u_j \rightharpoonup u$ in $W^{1,n-1}(\Omega, \mathbb{R}^{n-1})$, then we have $(\operatorname{cof}\nabla u_j)_i \rightharpoonup (\operatorname{cof}\nabla u)_i$ in $L^{\frac{n-1}{n-2}}(\Omega, \mathbb{R}^{n-1})$. Note that the i-th row of $(\operatorname{cof}\nabla v_j)$ is $(-1)^{i+k} \det \nabla_{\hat{x}_k} \hat{v}_j^i$ for $1 \le k \le n$ where $\nabla_{\hat{x}_k} \hat{v}_j^i$ is the $(n-1) \times (n-1)$ matrix of ∇v_j with ith row, kth column removed. For each ik-entry, we have $\det \nabla_{\hat{x}_k} \hat{v}_j^i = \frac{1}{n-1} \nabla_{\hat{x}_k} \hat{v}_j^i : \operatorname{cof}\nabla_{\hat{x}_k} \hat{v}_j^i$. By assumption, we know that $(\operatorname{cof}\nabla_{\hat{x}_k} \hat{v}_j^i)_m \rightharpoonup$

$(\mathrm{cof}\nabla_{\hat{x}_k}\hat{v}^i)_m$ in $L^{\frac{n-1}{n-2}}(\Omega,\mathbb{R}^{n-1})$ and $(\nabla_{\hat{x}_k}\hat{v}^i_j)_m \rightharpoonup (\nabla_{\hat{x}_k}\hat{v}^i)_m$ in $L^{n-1}(\Omega,\mathbb{R}^{n-1})$ for each row m. Thus, we have $\det\nabla_{\hat{x}_k}\hat{v}^i_j \overset{*}{\rightharpoonup} \det\nabla_{\hat{x}_k}\hat{v}^i$ in $C_0(\Omega)^*$. By the density of $C_0(\Omega)$ in $L^{\frac{n}{n-1}}(\Omega)$, we can deduce that $\sigma_j = (\mathrm{cof}\nabla v_j)_i \rightharpoonup (\mathrm{cof}\nabla v)_i$ in $L^{\frac{n}{n-1}}(\Omega,\mathbb{R}^n)$. Then the convergence follows from that

$$\frac{1}{n}\int_\Omega [(\nabla v_j)_i\cdot(\mathrm{cof}\nabla v_j)_i - (\nabla v)_i\cdot(\mathrm{cof}\nabla v)_i]\,\varphi\,\mathrm{d}x \to 0 \text{ for all } \varphi\in C_0(\Omega).$$

We can also use the Brezis and Nguyen (2011) theorem to show the result. $\square$

Let ν be an n-Young measure generated by a subsequence of $(\nabla v)_j$. Deduce that

$$\langle\nu_x,\det\rangle + \langle\nu_x^\infty,\det\rangle\,\frac{d\lambda}{d\mathcal{L}^n}(x) = \det\bar{\nu}_x \ \ \mathcal{L}^n a.e.$$

and

$$\langle\nu_x^\infty,\det\rangle = 0 \ \ \lambda^s a.e.$$

Proof. We know that $\nabla v_j \rightharpoonup \nabla v$ in $L^n(\Omega,\mathbb{R}^n)$, then $\nabla v(x) = \bar{\nu}_x$ *a.e.* From previous part of this question, we know that, for any $\varphi\in C_0(\Omega)$,

$$\int_\Omega \varphi\det\nabla v_j\,\mathrm{d}x \to \int_\Omega \varphi\det\bar{\nu}_x \text{ as } j\to\infty.$$

Since $\nu = ((\nu_x)_{x\in\Omega},\lambda,\nu_x^\infty)$ is the Young-measure generated by (∇v_j), then we have

$$\int_\Omega \varphi\det\nabla v_j\,\mathrm{d}x \to \int_\Omega \varphi\,\langle\nu_x,\det\rangle\,\mathrm{d}x + \int_{\bar{\Omega}}\varphi\,\langle\nu_x^\infty,\det\rangle\,\mathrm{d}\lambda \text{ as } j\to\infty.$$

This shows that

$$\int_\Omega \varphi\,\langle\nu_x,\det\rangle\,\mathrm{d}x + \int_{\bar{\Omega}}\varphi\,\langle\nu_x^\infty,\det\rangle\,\mathrm{d}\lambda = \int_\Omega \varphi\det\bar{\nu}_x \ \ a.e.$$

As φ is arbitrary, we have

$$\langle\nu_x,\det\rangle\,\mathrm{d}\mathcal{L}^n(x) + \langle\nu_x^\infty,\det\rangle\,\mathrm{d}\lambda(x) = \det\bar{\nu}_x\mathrm{d}\mathcal{L}^n(x).$$

Note that by decomposition, we have $\lambda = \lambda^s + \lambda^a$ for $\lambda^s,\lambda^a\in\mathcal{M}(\Omega)$ such that $\lambda^a \ll \mathcal{L}^n\lfloor\Omega$, $\lambda^s\perp\mathcal{L}^n\lfloor\Omega$. Then

$$\langle\nu_x,\det\rangle + \langle\nu_x^\infty,\det\rangle\,\frac{d\lambda^a}{d\mathcal{L}^n}(x) = \det\bar{\nu}_x \ \ \mathcal{L}^n a.e.$$

Thus

$$\langle\nu_x,\det\rangle + \langle\nu_x^\infty,\det\rangle\,\frac{d\lambda}{d\mathcal{L}^n}(x) = \det\bar{\nu}_x \ \ \mathcal{L}^n a.e.$$

Consider the set $A = \{x\in\Omega\colon \langle\nu_x,^\infty,\det\rangle \neq 0\}$. Then

$$\mathrm{d}\lambda(x) = \frac{(\det\bar{\nu}_x - \langle\nu_x,\det\rangle)\mathcal{L}^n(x)}{\langle\nu_x^\infty,\det\rangle} \text{ on } A.$$

This implies that $\mathrm{d}\lambda^a \ll \mathcal{L}^n\lfloor\Omega$ on A and $\mathrm{d}\lambda^s = 0$ on A. That is, $\langle\nu_x^\infty,\det\rangle = 0\ \lambda^s a.e.$ $\square$

Problem 2. *Let $Q\colon \mathbb{R}^{N\times n} \to \mathbb{R}$ be a quadratic form and let $g \in W^{1,2}(\Omega, \mathbb{R}^N)$. Consider the integral functional*

$$I[u] := \int_\Omega Q(\nabla u(x))\,\mathrm{d}x \;\; u \in W^{1,2}(\Omega, \mathbb{R}^N).$$

Show that $I[\cdot]$ is convex on the Dirichlet class $W^{1,2}_g(\Omega, \mathbb{R}^N)$ if and only if Q is rank-one convex. Give an example of a rank-one convex quadratic form for which the integral function $I[\cdot]$ is not convex on $W^{1,2}(\Omega, \mathbb{R}^N)$.

Proof. $I[\cdot]$ is convex on the Dirichlet class $W^{1,2}_g(\Omega, \mathbb{R}^N)$ if only if that for $u, v \in W^{1,2}_g(\Omega, \mathbb{R}^N)$, $\lambda \in (0, 1)$,

$$I[\lambda u + (1-\lambda)v] \le \lambda I[u] + (1-\lambda) I[v].$$

$$\Leftrightarrow \int_\Omega Q(\lambda\nabla u + (1-\lambda)\nabla v)\,\mathrm{d}x \le \lambda \int_\Omega Q(\nabla u)\,\mathrm{d}x + (1-\lambda)\int_\Omega Q(\nabla v)\,\mathrm{d}x.$$

$$\Leftrightarrow \int_\Omega \lambda^2 Q(\nabla u) + (1-\lambda)^2 Q(\nabla v) + 2\lambda(1-\lambda)\tilde{Q}\nabla u\cdot\nabla v\,\mathrm{d}x \le \lambda \int_\Omega Q(\nabla u)\,\mathrm{d}x + (1-\lambda)\int_\Omega Q(\nabla v)\,\mathrm{d}x.$$

$$\Leftrightarrow \int_\Omega 2\lambda(1-\lambda)\tilde{Q}\nabla u \cdot \nabla v\,\mathrm{d}x \le \int_\Omega \lambda(1-\lambda)Q(\nabla u) + \lambda(1-\lambda)Q(\nabla v)\,\mathrm{d}x.$$

$$\Leftrightarrow \int_\Omega 2\lambda(1-\lambda)\tilde{Q}\nabla u \cdot \nabla v\,\mathrm{d}x \le \int_\Omega \lambda(1-\lambda)\tilde{Q}\nabla u\cdot\nabla u + \lambda(1-\lambda)\tilde{Q}\nabla v\nabla v\,\mathrm{d}x.$$

$$\Leftrightarrow 0 \le \int_\Omega \lambda(1-\lambda)\tilde{Q}\nabla u \cdot (\nabla u - \nabla v)\,\mathrm{d}x - \int_\Omega \lambda(1-\lambda)\tilde{Q}\nabla v \cdot (\nabla u - \nabla v)\,\mathrm{d}x.$$

$$\Leftrightarrow 0 \le \lambda(1-\lambda)\int_\Omega Q(\nabla u - \nabla v)\,\mathrm{d}x$$

$$\Leftrightarrow 0 \le \frac{1}{|\Omega|}\int_\Omega Q(\nabla u - \nabla v)\,\mathrm{d}x \text{ since } \lambda \in (0,1).$$

$$\Leftrightarrow 0 \le \frac{1}{|\Omega|}\int_\Omega Q(\nabla \varphi)\,\mathrm{d}x \text{ for all } \varphi \in W^{1,2}_0(\Omega).$$

$$\overset{*}{\Leftrightarrow} Q(z) \le \frac{1}{|\Omega|}\int_\Omega Q(z + \nabla\varphi)\,\mathrm{d}x \text{ for all } z \in \mathbb{R}^{n\times N} \text{ and } \varphi \in W^{1,2}_0(\Omega).$$

$$\Leftrightarrow Q \text{ is quasiconvex.}$$

$$\Leftrightarrow Q \text{ is rank-one convex.}$$

Note that the implication indicated by $*$ is valid since

$$\begin{aligned}
\frac{1}{|\Omega|}\int_\Omega Q(z+\nabla\varphi)\,\mathrm{d}x =& \frac{1}{|\Omega|}\int_\Omega Q(z) + Q(\nabla\varphi) + 2\tilde{Q}z\cdot\nabla\varphi\,\mathrm{d}x \\
=& Q(z) + \frac{1}{|\Omega|}\int_\Omega Q(\nabla\varphi)\,\mathrm{d}x - \frac{2}{|\Omega|}\int_\Omega \nabla(\tilde{Q}z)\varphi\,\mathrm{d}x \\
=& Q(z) + \frac{1}{|\Omega|}\int_\Omega Q(\nabla\varphi)\,\mathrm{d}x \\
\ge& Q(z).
\end{aligned}$$

Counterexample: Let $Q := \det$. Consider $u, v \in W^{1,2}(\Omega, \mathbb{R}^2)$ defined as

$$u(x_1, x_2) = \begin{pmatrix} tx_2 \\ tx_1 \end{pmatrix} \text{ and } v(x_1, x_2) = \begin{pmatrix} sx_2 \\ sx_1 \end{pmatrix}$$

for $s \neq t \in \mathbb{R}$. From the lecture notes, we know that $z \mapsto \det z$ is quasi-affine, and thus rank-one convex. However, for $\lambda \in [0, 1]$,

$$\begin{aligned}
&I[\lambda u + (1-\lambda)v] - \lambda I[u] - (1-\lambda)I[v] \\
=&\int_\Omega \det(\nabla(\lambda u + (1-\lambda)v) - \lambda \int_\Omega \det(\nabla u)\,\mathrm{d}x - (1-\lambda)\int_\Omega \det(\nabla v)\,\mathrm{d}x \\
=&-(\lambda t + (1-\lambda)s)^2|\Omega| - (-\lambda t^2|\Omega|) - (-(1-\lambda)s^2|\Omega|) \\
=&|\Omega|\left(\lambda t^2 - (1-\lambda)s^2 - (\lambda t + (1-\lambda)s)^2\right) \\
=&|\Omega|\lambda(1-\lambda)(t-s)^2 \\
>&0 \text{ since } s \neq t.
\end{aligned}$$

That is,

$$I[\lambda u + (1-\lambda)v] > \lambda I[u] + (1-\lambda)I[v].$$

□

Problem 3. *Let $P\colon \mathbb{R}^{N\times n} \to \mathbb{R}$ be a rank-one convex polynomial of degree at most 3. Show that P must be quasiconvex.*

Proof. Since $P\colon \mathbb{R}^{N\times n} \to \mathbb{R}$ is a polynomial of degree at most 3, for $z \in \mathbb{R}^{N\times n}$,

$$P(z) = P(0) + P'(0)[z] + \frac{1}{2}P''(0)[z,z] + \frac{1}{6}P'''(0)[z,z,z]. \tag{0.1}$$

By rank-one convexity of P, we know that for $t \in \mathbb{R}$,

$$t \mapsto P(z_0 + ta \otimes b)$$

is convex for any $z_0 \in \mathbb{R}^{N\times n}$, $a \in \mathbb{R}^N$, $b \in \mathbb{R}^n$. Plugging $z = z_0 + ta \otimes b$ into (0.1), we have

$$\begin{aligned}
P(z_0 + ta\otimes b) =& P(0) + P'(0)[z_0 + ta\otimes b] + \frac{1}{2}P''(0)[z_0 + ta\otimes b, z_0 + ta\otimes b] \\
&+ \frac{1}{6}P'''(0)[z_0 + ta\otimes b, z_0 + ta\otimes b, z_0 + ta\times b] \\
=& P(0) + P'(0)[z_0] + \frac{1}{2}P''(0)[z_0, z_0] + \frac{1}{6}P'''(0)[z_0, z_0, z_0] \\
&+ t\left[P'(0)[a\otimes b] + P''(0)[z, a\otimes b] + \frac{1}{2}P'''(0)[z_0, z_0, a\otimes b\right] \\
&+ t^2\left[\frac{1}{2}P''(0)[a\otimes b, a\otimes b] + \frac{1}{2}P'''(0)[z_0, a\otimes b, a\otimes b]\right] \\
&+ t^3\left[\frac{1}{6}P'''(0)[a\otimes b, a\otimes b, a\otimes b]\right].
\end{aligned}$$

Since $t \mapsto P(z_0 + ta \otimes b)$ is convex, we can deduce that

$$P'''(0)[a \otimes b, a \otimes b, a \otimes b] = 0 \text{ for all } a \in \mathbb{R}^N, b \in \mathbb{R}^n. \tag{0.2}$$

and

$$P''(0)[a \otimes b, a \otimes b] + P'''(0)[z_0, a \otimes b, a \otimes b] \geq 0 \text{ for all } z \in \mathbb{R}^{N\times n}, a \in \mathbb{R}^N, b \in \mathbb{R}^n. \tag{0.3}$$

(0.2) and (0.3) implies that

$$P''(0)[a \otimes b, a \otimes b] \geq 0 \text{ for all } a \in \mathbb{R}^N, b \in \mathbb{R}^n. \tag{0.4}$$

Note that since (0.3) works for all $z_0 \in \mathbb{R}^{N\times n}$, we must have

$$P'''(0)[z_0, a \otimes b, a \otimes b] = 0 \text{ for all } a \in \mathbb{R}^N, b \in \mathbb{R}^n. \tag{0.5}$$

Note that the first two terms $P(0) + P'(0)[z]$ in (0.1) is affine, thus quasi-convex. For the 2nd term $\frac{1}{2}P''(0)[z, z]$, first note that it is quadratic form.

$$P''(0)[z_0 + ta \otimes b, z_0 + ta \otimes b] = P''(0)[z_0, z_0] + 2tP''(0)[z_0, a \otimes b] + t^2 P''(0)[a \otimes b, a \otimes b].$$

Note that $\frac{d^2}{dt^2}P''(0)[z_0+ta\otimes b, z_0+ta\otimes b] = P''(0)[a\otimes b, a\otimes b] \geq 0$ by (0.4), then $P''(0)[z, z]$ is rank-one convex by the Legendre-Hadamard condition. Then using the result from Q2 for quadratic forms, we know that the second term $\frac{1}{2}P''(0)[z, z]$ is quasi-convex as well. It remains to show the last term $\frac{1}{6}P'''(0)[z, z, z]$ is quasi-convex. Now we aim to show that $P'''(0)[z, z, z]$ is rank-one affine, then by a theorem from Ball which states that rank-one affine and quasi-affine are equivalent, we can conclude that $P'''(0)[z, z, z]$ is quasi-affine, and thus quasi-convex. Note that

$$\begin{aligned}
&P'''(0)[z_0 + ta \otimes b, z_0 + ta \otimes b, z_0 + ta \otimes b] \\
=&P'''(0)[z_0, z_0, z_0] + 3tP'''(0)[z_0, z_0, a \otimes b] + 3t^2 P'''(0)[z_0, a \otimes b, a \otimes b] + t^3 P^{'''}(0)[a \otimes b, a \otimes b, \otimes b] \\
=&P'''(0)[z_0, z_0, z_0] + 3tP'''(0)[z_0, z_0, a \otimes b] + 3t^2 P'''(0)[z_0, a \otimes b, a \otimes b] \text{ (by (0.2))} \\
=&P'''(0)[z_0, z_0, z_0] + 3tP'''(0)[z_0, z_0, a \otimes b] \text{ (by (0.5))}
\end{aligned}$$

is rank-one affine.
By subtracting an affine function, we can without loss of generality assume that $P = P_2 + P_3$ where degree of P_i is i and P_i is homogeneous for $I = 2, 3$. Note that

$$P_3(A) = \lim_{t\to\infty} \frac{P(tA)}{t^3}.$$

This shows that P_3 is the limit of rank-one convex functions, and thus rank-one convex. Since P_3 is odd and it is rank-one convex, then it is rank-affine. Then $P_2 = P - P_3$ is rank-one convex. But P_2 is quadratic, so it is quasi-convex. □

Problem 4. *(a) Let $f\colon \mathbb{R} \to \mathbb{R}$ be a convex function. Show the 3-slope inequality:*

$$\frac{f(s)-f(r)}{s-r} \leq \frac{f(t)-f(r)}{t-r} \leq \frac{f(t)-f(s)}{t-s}$$

for all $r < s < t$. Deduce that f is differentiable except for at most countably many points and that

$$\operatorname*{ess\,sup}_{s\in(-r,r)} |f'(s)| \leq \frac{2}{r} \max_{s\in[-2r,2r]} |f(s)|.$$

Proof. Since f is convex, for $\lambda \in (0,1)$, we have

$$f(\lambda r + (1-\lambda)t) \leq \lambda f(r) + (1-\lambda) f(t)$$

for all $r < t$. Take $\lambda = \frac{t-s}{t-r}$ and $1-\lambda = \frac{s-r}{t-r}$, then $\lambda r + (1-\lambda)t = s$. Thus

$$f(s) = f(\lambda r + (1-\lambda)t) \leq \frac{t-s}{t-r} f(r) + \frac{s-r}{t-r} f(t).$$

Subtracting $f(r)$ on both sides and dividing by $s-r$, we have

$$\frac{f(s)-f(r)}{s-r} \leq \frac{f(t)-f(r)}{t-r}$$

for $r < s < t$.

Now if we subtract $f(t)$ on both sides and then divide by $t-s$, we have

$$\frac{f(s)-f(t)}{t-s} \leq \frac{f(r)-f(t)}{t-r}.$$

Multiplying by (-1) on both sides, we have

$$\frac{f(t)-f(s)}{t-s} \geq \frac{f(t)-f(r)}{t-r}.$$

From the 3-slope inequality, we can deduce that f has one-sided derivatives at each point and that $f'_-(x) \leq f'_+(x) \leq f'_-(y) \leq f'_+(y)$ for each $x < y$. For every point, where $f'_-(x) < f'_+(x)$, we can chose a rational number q such that $f'_-(x) < q < f'_+(x)$, and these rational numbers are distinct. Thus, f is differentiable except for at at most countably many points.

Consider $-2r < -r < s < r < 2r$, then the 3-slope inequality implies that for $-r < s < \tilde{s} < r < 2r$,

$$f'_+(s) = \lim_{\tilde{s}\to s} \frac{f(\tilde{s})-f(s)}{\tilde{s}-s} < \frac{f(2r)-f(s)}{2r-s} \leq \frac{f(2r)-f(s)}{2r-r} \leq \frac{2}{r} \max_{s\in[-2r,2r]} |f(s)|.$$

Similarly, for $-r < s' < s < r < 2r$,

$$f'_-(s) = \lim_{s'\to s} \frac{f(s)-f(s')}{s-s'} < \frac{f(2r)-f(s)}{2r-s} \leq \frac{f(2r)-f(s)}{2r-r} \leq \frac{2}{r} \max_{s\in[-2r,2r]} |f(s)|.$$

Thus,

$$\operatorname*{ess\ sup}_{s\in(-r,r)} |f'(s)| \le \frac{2}{r} \max_{s\in[-2r,2r]} |f(s)|.$$

$\square$

(b) Assume that $F\colon \mathbb{R}^{N\times n} \to \mathbb{R}$ *is rank-one convex and differentiable. Denote by* $\|\cdot\|$ *the maximum norm on* $\mathbb{R}^{N\times n}$*;*

$$\|z\| := \max\{|z_{r,s}|\colon 1 \le r \le N, 1 \le s \le n\}.$$

Show that

$$\sup_{\|z\|\le r} \|F'(z)\| \le \frac{2}{r} \max_{\|z\|\le 2r} |F(z)|$$

holds for all $r > 0$*. Here we define* $F'(z) \in \mathbb{R}^{N\times n}$ *by*

$$F'(z)\cdot w := \frac{d}{dt}\,|_{t=0}\ F(z+tw)\ \ \forall w \in \mathbb{R}^{N\times n}.$$

Proof. Since F is rank-one convex and differentiable, $t \mapsto F(z + ta\otimes b)$ is convex. By (a), we know that

$$\operatorname*{ess\ sup}_{t\in(-r,r)} \left|\frac{d}{dt} F(z+ta\otimes b)\right| \le \frac{2}{r} \max_{t\in[-2r,2r]} |F(z+ta\otimes b)|.$$

This implies that

$$\max_{\|z\|<r} |F'(z)\cdot(a\otimes b)| \le \frac{2}{r} \max_{\|z\|<r, t\in[-2r,2r]} |F(z+ta\otimes b)|$$

for all $a \in \mathbb{R}^N$, $b \in \mathbb{R}^n$. Taking $a = e_i = [0,\dots,1,\dots,0]^T$ where 1 appears at the ith entry and the rest are zeros, $b = e_j$ for each $1 \le i \le N$, $1 \le j \le n$. Then we can conclude that

$$\max_{\|z\|<r} \|F'(z)\| \le \frac{2}{r} \max_{\|z\|\le 2r} |F(z)|.$$

$\square$

Using a standard smooth mollifier show that a rank-one convex integrand (possibly non-differentiable) $G\colon \mathbb{R}^{N\times n} \to \mathbb{R}$ *is locally Lipschitz. Show also that there exists* $z_0 \in \mathbb{R}^{N\times n}$ *such that*

$$G(z) \ge G(0) + z_0\cdot z$$

holds for all $z \in \mathbb{R}^{N\times n}$ *with* $\operatorname{rank}(z) = 1$.

Proof. We first show that G is continuous at z. Without loss of generality, we may assume that $z = 0$ and $G(0) = 0$. Since a real-valued rank-one convex integrand is locally bounded, G is bounded above in a neighbourhood of $z = 0$. That is, there exists $\delta > 0$ and $a > 0$ such that

$$\|z\| \leq \delta, \Rightarrow G(z) \leq a.$$

For all $\varepsilon > 0$ we can find $\delta > 0$ such that

$$\|z\| < \frac{\varepsilon}{aN2^N}\delta \Rightarrow |G(z)| \leq \varepsilon.$$

Without loss of generality, we can assume that $\varepsilon < aN2^N$ or otherwise we can simply increase the upper bound a, and define $\lambda := \frac{\varepsilon}{aN2^N}$. Using the rank-one convexity of G, we have

$$\begin{aligned}
G(z) =& G(\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_N \end{bmatrix}) \text{ (where} z_i, 1 \leq i \leq n \text{ is the ith row of } z) \\
=& G\left(\lambda \begin{bmatrix} z_1/\lambda \\ z_2 \\ \vdots \\ z_N \end{bmatrix} + (1-\lambda) \begin{bmatrix} 0 \\ z_2 \\ \vdots \\ z_N \end{bmatrix}\right) \\
\leq& \lambda G\left(\begin{bmatrix} z_1/\lambda \\ z_2 \\ \vdots \\ z_N \end{bmatrix}\right) + (1-\lambda) G\left(\begin{bmatrix} 0 \\ z_2 \\ \vdots \\ z_N \end{bmatrix}\right).
\end{aligned}$$

Repeating the process with the second row, we have

$$G(z) \leq \lambda G\left(\begin{bmatrix} z_1/\lambda \\ z_2 \\ z_3 \\ \vdots \\ z_N \end{bmatrix}\right) + (1-\lambda)\lambda G\left(\begin{bmatrix} 0 \\ z_2/\lambda \\ z_3 \\ \vdots \\ z_N \end{bmatrix}\right) + (1-\lambda)^2 G\left(\begin{bmatrix} 0 \\ 0 \\ z_3 \\ \vdots \\ z_N \end{bmatrix}\right).$$

Iterating this process, we obtain

$$G(z) \leq \lambda \sum_{i=1}^{N} (1-\lambda)^{i-1} G\left(\begin{bmatrix} 0 \\ \vdots \\ z_i/\lambda \\ \vdots \\ z_N \end{bmatrix}\right) + (1-\lambda)^N G(0).$$

Now we assume that $\|z\|/\lambda \le \delta$, that is, $\|z\| \le \lambda\delta \le \delta$, so $G(z) < a$, and

$$G(z) \le \lambda \sum_{i=1}^{N} (1-\lambda)^{i-1} a \le \lambda a N < \frac{\varepsilon}{2^N} < \varepsilon.$$

In order to prove the lower bound, we proceed the proof in a similar manner.

$$\begin{aligned} 0 =& G(0) \\ =& G\left(\frac{1}{\lambda+1} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ z_N \end{bmatrix} + \frac{\lambda}{1+\lambda} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ -z_N/\lambda \end{bmatrix} \right) \\ \le& \frac{1}{\lambda+1} G\left(\begin{bmatrix} 0 \\ 0 \\ \vdots \\ z_N \end{bmatrix} \right) + \frac{\lambda}{1+\lambda} G\left(\begin{bmatrix} 0 \\ 0 \\ \vdots \\ -z_N/\lambda \end{bmatrix} \right). \end{aligned}$$

Repeating with the $(N-1)$th row,we get

$$0 \le \frac{1}{(1+\lambda)^2} G\left(\begin{bmatrix} 0 \\ \vdots \\ z_{N-1} \\ z_N \end{bmatrix} \right) + \frac{\lambda}{(1+\lambda)^2} G\left(\begin{bmatrix} 0 \\ \vdots \\ -z_{N-1}/\lambda \\ z_N \end{bmatrix} \right) + \frac{\lambda}{1+\lambda} G\left(\begin{bmatrix} 0 \\ \vdots \\ 0 \\ -z_N/\lambda \end{bmatrix} \right).$$

Iterating the process above, we deduce that

$$0 \le \frac{1}{(1+\lambda)^N} G(z) + \sum_{i=1}^{N} \frac{\lambda}{(1+\lambda)^{N-i+1}} G\left(\begin{bmatrix} 0 \\ \vdots \\ z_i/\lambda \\ z_{i+1} \\ \vdots \\ z_N \end{bmatrix} \right).$$

Hence if $\|z\|/\lambda \leq \delta$, that is, $\|z\| \leq \lambda\delta \leq \delta$, we have

$$
\begin{aligned}
G(z) \geq & -\lambda \sum_{i=1}^{N} (1+\lambda)^{i-1} G\left(\begin{bmatrix} 0 \\ \vdots \\ z_i/\lambda \\ z_{i+1} \\ \vdots \\ z_N \end{bmatrix}\right) \\
\geq & -\lambda a \sum_{i=1}^{N} (1+\lambda)^{i-1} \\
\geq & -\lambda a N 2^N \\
= & -\varepsilon.
\end{aligned}
$$

This shows that G is continuous at $z = 0$. Thus G is continuous at z for any $z \in \mathbb{R}^{N\times n}$. Now consider the standard mollifier ρ defined as

$$
\rho(z) = \begin{cases} C \exp\left(\frac{1}{\|z\|-1}\right) & \text{if } \|z\| < 1, \\ 0 & \text{otherwise.} \end{cases}
$$

where C is chosen such that $\int_{\mathbb{R}^{N\times n}} \rho(z)\,\mathrm{d}z = 1$. $\rho_\varepsilon(z) = \frac{1}{\varepsilon^n}\rho(\frac{z}{\varepsilon})$. Consider

$$
F(z) := \rho_\varepsilon * G = \int_{\mathbb{R}^{N\times n}} \rho_\varepsilon(z-y) G(y)\,\mathrm{d}y.
$$

First note that F is smooth with compact support and F is still rank-one convex. Let $z_1, z_0 \in \mathbb{R}^{N\times n}$, $\operatorname{rank}(z_1 - z_0) = 1$, and for $\lambda \in (0,1)$, we know that

$$
G(\lambda z_1 + (1-\lambda) z_0) \leq \lambda G(z_1) + (1-\lambda) G(z_0).
$$

Note that

$$
F(z) = \int_{\mathbb{R}^{N\times n}} \rho_\varepsilon(z-y) G(y)\,\mathrm{d}y = \int_{\mathbb{R}^{N\times n}} \rho_\varepsilon(y) G(z-y)\,\mathrm{d}y
$$

by a change of variable and symmetry of the mollifier. Then

$$
\begin{aligned}
F(\lambda z_1 + (1-\lambda) z_0) = & \int_{\mathbb{R}^{N\times n}} \rho_\varepsilon(y) G(\lambda(z_1) + (1-\lambda) z_0 - y)\,\mathrm{d}y \\
= & \int_{\mathbb{R}^{N\times n}} \rho_\varepsilon(y) G(\lambda(z_1 - y) + (1-\lambda)(z_0 - y))\,\mathrm{d}y \\
\leq & \int_{\mathbb{R}^{N\times n}} \rho_\varepsilon(y) \left[\lambda G(z_1 - y) + (1-\lambda) G(z_0 - y)\right]\,\mathrm{d}y \\
\leq & \lambda F(z_1) + (1-\lambda) F(z_0)
\end{aligned}
$$

since $z_1 - y$ and $z_0 - y$ are rank-one connected. Also, we know that F converges to G locally uniformly. That is, for $z \in \mathbb{R}^{N\times n}$,

$$|F(z) - G(z)| \to 0 \text{ as } \varepsilon \to 0.$$

Then for $z_1, z_0 \in \mathbb{R}^{N\times n}$,

$$\begin{aligned}
|G(z_1) - G(z_0)| \leq & |F(z_1) - G(z_1)| + |F(z_0) - G(z_0)| + |F(z_1) - F(z_0)| \\
\leq & |F(z_1) - G(z_1)| + |F(z_0) - G(z_0)| + \sup_{\|z\|\leq r} \|F'(z)\| \|z_1 - z_0\| \\
\leq & |F(z_1) - G(z_1)| + |F(z_0) - G(z_0)| + \frac{2}{r} \max_{\|z\|\leq 2r} |F(z)| \|z_1 - z_0\| \\
\leq & |F(z_1) - G(z_1)| + |F(z_0) - G(z_0)| + \frac{2}{r} \max_{\|z\|\leq 2r} |G(z)| \|z_1 - z_0\| \\
\to & \frac{2}{r} \max_{\|z\|\leq 2r} |G(z)| \|z_1 - z_0\| \text{ as } \varepsilon \to 0.
\end{aligned}$$

where z in the inequality is in between z_0 and z_1. Note that G is locally bounded, then from the above inequality we can conclude that G is locally Lipschitz continuous.

For each $z \in \mathbb{R}^{N\times n}$, we have

$$F(z) = F(0) + F'(0)[z] + \frac{1}{2} F''(\xi)[z, z] \text{ for some } \xi \text{ in between } 0, z.$$

Since F is rank-one convex, we know that $t \mapsto F(z_0 + ta \otimes b)$ is convex for all $z_0 \in \mathbb{R}^{N\times n}$, $a \in \mathbb{R}^N$ and $b \in \mathbb{R}^n$. By convexity, we have $F''(z_0)[z, z] \geq 0$ for any $z_0 \in \mathbb{R}^{N\times n}$. Then we know that

$$F(z) \geq F(0) + F'(0)[z].$$

We know that $F(z) \to G(z)$ as $\varepsilon \to 0$, $F(0) \to G(0)$ as $\varepsilon \to 0$. We know that

$$\|F'(0)\| \leq \frac{2}{r} \max_{\|z\|\leq 2r} |F(z)| \leq \frac{2}{r} \max_{\|z\|<r} |G(z)| < \infty$$

since G is locally bounded. Then there exists a convergent subsequence, also call it $F'(0)$ such that $F'(0) \to z_0$ as $\varepsilon \to 0$.

So we can conclude that

$$G(z) \geq G(0) + z_0 \cdot z$$

for some $z_0 \in \mathbb{R}^{N\times n}$.

$\square$